118892

From: Whiteman, Brian
Sent: Tuesday, April 06, 2004 6:34 PM
To: STIC-Biotech/ChemLib
Subject: seq search

09/610,313
7/5/00
Barnett et al.,

search SEQ ID NO: 30, 31, and 32 against us patent and us patent application databases.

Thanks,
Brian Whiteman
Remsen, 2D14
mail box 2C18
Patent Examiner - Art Unit 1635
United States Patent and Trademark Office
(571) 272-0764

Searcher:_____
Phone:_____
Location:_____
Date Picked Up:_____
Date Completed:_____
Searcher Prep/Review:_____
Clerical:_____
Online time:_____

TYPE OF SEARCH:_____
NA Sequences:_____3_____
AA Sequences:_____
Structures:_____
Bibliographic:_____
Litigation:_____
Full text:_____
Patent Family:_____
Other:_____

VENDOR/COST (where applic.)
STN:_____
DIALOG:_____
Questel/Orbit:_____
DRLink:_____
  Lexis/Nexis:_____
Sequence Sys.:_____
WWW/Internet:_____
  Other (specify):_____

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.
GenCore version 5.1.6

Run on: April 10, 2004, 02:53:16 ; Search time 6311.25 Seconds
                  (without alignments)
                  16873.640 Million cell updates/sec

Title: US-09-610-313-32
Perfect score: 2457
Sequence: 1 gtcgacgccacatggccga..........gggctagcaccggtgaattc 2457

Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45 summaries

Database :   GenEmbl:*
                  1: gb_ba:*
                  2: gb_htg:*
                  3: gb_in:*
                  4: gb_om:*
                  5: gb_ov:*
                  6: gb_pat:*
                  7: gb_ph:*
                  8: gb_pl:*
                  9: gb_pr:*
                  10: gb_ro:*
                  11: gb_sts:*
                  12: gb_sy:*
                  13: gb_un:*
                  14: gb_vi:*
                  15: gb_vr:*
                  16: em_fun:*
                  17: em_hum:*
                  18: em_in:*
                  19: em_mu:*
                  20: em_om:*
                  21: em_or:*
                  22: em_ov:*
                  23: em_pat:*
                  24: em_ph:*
                  25: em_pl:*
                  26: em_ro:*
                  27: em_sts:*
                  28: em_un:*
                  29: em_vi:*
                  30: em_hum:*
                  31: em_htg_inv:*
                  32: em_htg_other:*
                  33: em_htg_mus:*
                  34: em_htg_pln:*
                  35: em_htg_rod:*
                  36: em_htg_mam:*
                  37: em_htg_vrt:*
                  38: em_sy:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2457	100.0	2457	6	AX455916	AX455916 Sequence
2	2436.2	99.2	2463	6	AX455915	AX455915 Sequence
3	2415.4	98.3	2469	6	AX455914	AX455914 Sequence
4	2040	83.0	2300	6	BD263705	BD263705 Improved
5	2040	83.0	2300	6	BD263704	BD263704 Sequence
6	2019.2	82.2	2306	6	BD263704	BD263704 Improved
7	2019.2	82.2	2306	6	AR373387	AR373387 Sequence
8	2000.8	81.4	2312	6	AX427930	AX427930 Sequence
9	1998.4	81.3	2312	6	BD263706	BD263706 Improved
10	1998.4	81.3	2312	6	AR373389	AR373389 Sequence
11	1992.4	81.1	2312	6	AX427931	AX427931 Sequence
12	1978.4	80.5	9788	6	AX427936	AX427936 Sequence
13	1975	80.4	9169	6	AX427926	AX427926 Sequence
14	1973.4	80.3	9194	6	AX427925	AX427925 Sequence
15	1970	80.2	12411	6	AX427927	AX427927 Sequence
16	1966.6	80.0	9785	6	AX427938	AX427938 Sequence
17	1966.6	80.0	9167	6	AX427933	AX427933 Sequence
18	1966.6	80.0	9170	6	AX427928	AX427928 Sequence
19	1966.6	80.0	9783	6	AX427935	AX427935 Sequence
20	1966.6	80.0	9782	6	AX427934	AX427934 Sequence
21	1966.6	80.0	9792	6	AX427932	AX427932 Sequence
22	1965	80.0	9189	12	AX427921	AX427921 Sequence
23	1962	79.9	4353	12	AF287353	AF287353 Synthetic
24	1945.4	79.2	3015	6	AX455946	AX455946 Sequence
25	1928.8	78.5	3009	6	AX455987	AX455987 Sequence
26	1915.8	78.0	4319	6	BD263639	BD263639 Improved
27	1915.8	78.0	4319	6	AR373322	AR373322 Sequence
28	1905.2	77.5	4352	12	AF287352	AF287352 Synthetic
29	1878.6	76.5	1965	6	AX455952	AX455952 Sequence
30	1872.8	76.2	2299	6	BD263703	BD263703 Improved
31	1872.8	76.2	2299	6	AR373386	AR373386 Sequence
32	1859.8	75.7	2577	6	AX457088	AX457088 Sequence
33	1852.2	75.4	4343	6	AX306428	AX306428 Sequence
34	1852.2	75.4	2305	6	BD263702	BD263702 Improved
35	1852	75.4	2305	6	AR373385	AR373385 Sequence
36	1848.2	75.2	4341	6	AX306429	AX306429 Sequence
37	1838.6	74.8	1977	6	AX455954	AX455954 Sequence
38	1825.6	74.3	7897	6	AX427923	AX427923 Sequence
39	1772	72.1	3024	12	AF287355	AF287355 Synthetic
40	1757.2	71.5	9407	6	AX427937	AX427937 Sequence
41	1626.6	66.2	4307	6	AR302563	AR302563 Sequence
42	1626.6	66.2	4307	6	AX019132	AX019132 Sequence
43	1626.6	66.2	4307	6	BD130468	BD130468 Antiviral
44	1607.4	65.4	4307	6	BD268901	BD268901 Anti-vira
45	1607.4	65.4	4307	6	AX035453	AX035453 Sequence

ALIGNMENTS

RESULT 1
AX455916
LOCUS   AX455916   2457 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION Sequence 32 from Patent WO0204493.
ACCESSION   AX455916
VERSION   AX455916.1 GI:21714901
KEYWORDS
SOURCE   synthetic construct
ORGANISM   synthetic construct
            artificial sequences.

```
FEATURES
    source          CHIRON CORPORATION (US); University of Stellenbosch (ZA)
                    Location/Qualifiers
                    1..2457
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
                    /note="PR975MWM"

ORIGIN

Query Match       100.0%;  Score 2457;  DB 6;  Length 2457;
Best Local Similarity  100.0%;  Pred. No. 2.4e-248;
Matches 2457;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 GTCGACGCGCCACCATGGCCAGGCCATGAGCCAGGCGCCAACATCCTGATGCAG    60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GTCGACGCGCCACCATGGCCAGGCCATGAGCCAGGCGCCAACATCCTGATGCAG    60

Qy   61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGCTTCAACTCGGCAAGAGGGC   120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGCTTCAACTCGGCAAGAGGGC   120

Qy  121 CACATCGCCCGCAACTGCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGCAAGGAG   180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CACATCGCCCGCAACTGCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGCAAGGAG   180

Qy  181 GGCCACCAGATGAAGGACTGCACCGAGGCGCCAGCCAACTTCTTCCCGCGAGGACCTGGCC  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GGCCACCAGATGAAGGACTGCACCGAGGCGCCAGCCAACTTCTTCCCGCGAGGACCTGGCC  240

Qy  241 TTCCCCAGGGCAAGGCCCCGGCGAGTTCCCAGCGAGCAGAACCGGCCGCGGCCGAGCGCCACC  300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TTCCCCAGGGCAAGGCCCCGGCGAGTTCCCAGCGAGCAGAACCGGCCGCGGCCGAGCGCCACC  300

Qy  301 AGCCGCGAGTCGCAGGTGCGCGGGCGACAACCCCGCGAGCGAGGCGGCCGGAGCGGCCAG   360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AGCCGCGAGTCGCAGGTGCGCGGGCGACAACCCCGCGAGCGAGGCGGCCGGAGCGGCCAG   360

Qy  361 GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG   420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG   420

Qy  421 GGCGGCCAGATCAAGGAGGAGGCCCTGCTGACACCGGCGCGCGCCGGTGCTGCTGGAGGAG  480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGCGGCCAGATCAAGGAGGAGGCCCTGCTGACACCGGCGCGCGCCGGTGCTGCTGGAGGAG  480

Qy  481 ATGAGCCTGCCCCGGCAAGTGGAAGCCCAAGTGCCGTGCCCATCAGCGCCGCTTCATCAAG   540
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 ATGAGCCTGCCCCGGCAAGTGGAAGCCCAAGTGCCGTGCCCATCAGCGCCGCTTCATCAAG   540

Qy  541 GTGGCCCAGTACGACCAGATCCTGATCGAGATCGCGCAAGAAGGCCATCGGCACCGTG   600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 GTGGCCCAGTACGACCAGATCCTGATCGAGATCGCGCAAGAAGGCCATCGGCACCGTG   600

Qy  601 CTGATCGGCCCCACCCCCTGAACATCGGCGCGACATCATCGGCGGCTGGGCTGC   660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 CTGATCGGCCCCACCCCCTGAACATCGGCGCGACATCATCGGCGGCTGGGCTGC   660

Qy  661 ACCCTGAACTTCCCCATCAGCCCATGCCCGTGAAGCTGAAGCCCCGGCATG   720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661 ACCCTGAACTTCCCCATCAGCCCATGCCCGTGAAGCTGAAGCCCCGGCATG   720

Qy  721 GACGGCCCAAGGTGAAGCAGTGCCCCTGACCAAGGCCCTGACCGCC   780
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 GACGGCCCAAGGTGAAGCAGTGCCCCTGACCAAGGCCCTGACCGCC   780

Qy  781 ATCTGCGAGGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC   840
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 ATCTGCGAGGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC   840

Qy  841 AACACCCCGTGTTCGCCATCAAGAAGAGACACCAAGTGGCGCAAGTCGTGGAC   900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841 AACACCCCGTGTTCGCCATCAAGAAGAGACACCAAGTGGCGCAAGTCGTGGAC   900
```

```
Qy  901 TTCCGCGGAGCTGAACAAGGCGACCCAGGACTTCTCGGAGGTGCAGCTGGCATCCCCAC   960
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTCCGCGGAGCTGAACAAGGCGACCCAGGACTTCTCGGAGGTGCAGCTGGCATCCCCAC   960

Qy  961 CCGCGCGGCCTGAAGAGAAGAGAGAGGGTGACCGTGCTGGACGTGGCGACGCTGGCACGCGCCTACTTC  1020
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 CCGCGCGGCCTGAAGAGAAGAGAGAGGGTGACCGTGCTGGACGTGGCGACGCTGGCACGCGCCTACTTC  1020

Qy 1021 AGGGTGCCCTGGACGAGGACTTCCGAAGTACACCGCCTTCACCACCATCCCAGCATCAAC  1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 AGGGTGCCCTGGACGAGGACTTCCGAAGTACACCGCCTTCACCACCATCCCAGCATCAAC  1080

Qy 1081 AACGAGACCCCCGCATCGTCTTGGTACCAGTACAACGTGTCGCCCAGGGCTGGAAGGGCAGC  1140
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 AACGAGACCCCCGCATCGTCTTGGTACCAGTACAACGTGTCGCCCAGGGCTGGAAGGGCAGC  1140

Qy 1141 CCAGCAATCTTCAGAGCAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCGCAACCCC  1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 CCAGCAATCTTCAGAGCAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCGCAACCCC  1200

Qy 1201 GAGATCGTGATCTACCAGCCCCGTGACGTGGGCAGCAGCGACCTGGCAGCAC  1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GAGATCGTGATCTACCAGCCCCGTGACGTGGGCAGCAGCGACCTGGCAGCAC  1260

Qy 1261 CGGCCAAGATCGAGAGACTGCGCAAGCACCTGCTGCCGCTTCACCACCCCCGAC  1320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 CGGCCAAGATCGAGAGACTGCGCAAGCACCTGCTGCCGCTTCACCACCCCCGAC  1320

Qy 1321 AAGAAGCACCAGAAGGAGCCCCCCTTCTGCCCATGCAGCTGACCGTGAACCGACATCCAGAAGCTG  1380
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 AAGAAGCACCAGAAGGAGCCCCCCTTCTGCCCATGCAGCTGACCGTGAACCGACATCCAGAAGCTG  1380

Qy 1381 GTGCAGCCCATGGACTGCCCGAGAAGGAGCTGCCCGAGCGAGTCTACCCCGGCATCATCCAGACCTACGAC  1440
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381 GTGCAGCCCATGGACTGCCCGAGAAGGAGCTGCCCGAGCGAGTCTACCCCGGCATCATCCAGACCTACGAC  1440

Qy 1441 GTGGGCAAGCTGAACTCGGCCAGCCAGATCTACCCCGGCATCATCAAGGTGCCCAGCTGTGC  1500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 GTGGGCAAGCTGAACTCGGCCAGCCAGATCTACCCCGGCATCATCAAGGTGCCCAGCTGTGC  1500

Qy 1501 AAGCTGCTGCGGGGCCGCCAAGGCCCCTGACCGACATCGTGCCCGACGAGGAGCCGAG  1560
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 AAGCTGCTGCGGGGCCGCCAAGGCCCCTGACCGACATCGTGCCCGACGAGGAGCCGAG  1560

Qy 1561 CTTGGAGCTGGCCGAGAACCGGCAGATCTTGCGCGAGCCGTGCACGGCGTGTACTACGAC  1620
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 CTTGGAGCTGGCCGAGAACCGGCAGATCTTGCGCGAGCCGTGCACGGCGTGTACTACGAC  1620

Qy 1621 CCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGCAGGGCCACGACCAGACCTACCAG  1680
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 CCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGCAGGGCCACGACCAGACCTACCAG  1680

Qy 1681 ATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGATGCCGACCGGCC  1740
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 ATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGATGCCGACCGGCC  1740

Qy 1741 CACCAACGACGTGAAGCAGCTGACCGAGCCGTGCAGAGATCGCCATGGAGAGCATC  1800
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 CACCAACGACGTGAAGCAGCTGACCGAGCCGTGCAGAGATCGCCATGGAGAGCATC  1800

Qy 1801 GTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCAGGAGACCTGGGAGACC  1860
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 GTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCAGGAGACCTGGGAGACC  1860

Qy 1861 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGGAGTTCGTGAAACCCCCC  1920
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGGAGTTCGTGAAACCCCCC  1920

Qy 1921 CCCCTGGTGAAGCTGTGGTACCAGTGCGCAAGGCACCAGTGCCAAGCTTC  1980
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 CCCCTGGTGAAGCTGTGGTACCAGTGCGCAAGGCACCAGTGCCAAGACTTC  1980

Qy 1981 TACGTGACGCGGCGCCAACCGCGAGACATCGGCAGAGGCCGCTCACGTGACCGAC  2040
```

```
Db   1981  TACGTGGACGGCGCGCCCACCCGAGACCAAGATCGGCAAGACCAGCCAACTTCTTCCGGAGGACCTGACCGAC  2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2041  CGGGGCCGGCGAAGATCGTGAGCCTGAGCGCCGAGACCAACCAACCAGAAGACCGAGCTGCAG  2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041  CGGGGCCGGCGAAGATCGTGAGCCTGAGCGCCGAGACCAACCAACCAGAAGACCGAGCTGCAG  2100

Qy   2101  GCCATCCAGTGGCCCTGCAGGACAGCGGAGGTGAACATCGTGACGACAGCCAG  2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  GCCATCCAGTGGCCCTGCAGGACAGCGGAGGTGAACATCGTGACGACAGCCAG  2160

Qy   2161  TACGCCCTGGGCATCATCCAGGACGGCAGCGGCGCGGAGAGACCAGAGCTGGTGAACCAG  2220
            |||||||||||||||||(||)||||||||||||||||||||||||||||||||||||||
Db   2161  TACGCCCTGAACTTCCCCGGCATCATCCAGGACGGCAGCGGCGCGGAGAGACCAGAGCTGGTGAACCAG  2220

Qy   2221  ATCATCGAGCAGCTGATCAAGAAGAGAGGTGTACCTGAGCTGGTGGCAAG  2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  ATCATCGAGCAGCTGATCAAGAAGAGAGGTGTACCTGAGCTGGTGGCAAG  2280

Qy   2281  GGCATGGCGCAACGAGCAGATCGACAAGTGGTGAGCAAGGGCATCCGGCAAGTGCTG  2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  GGCATGGCGCAACGAGCAGATCGACAAGTGGTGAGCAAGGGCATCCGGCAAGTGCTG  2340

Qy   2341  TTCCTGACGGGATGATGGCGCATCGTGATCTACCAGTACATGGCGACGACCTGTACGTG  2400
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  TTCCTGACGGGATGATGGCGCATCGTGATCTACCAGTACATGGCGACGACCTGTACGTG  2400

Qy   2401  GGCAGCGGGCGGCGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCCGGTGAATTC  2457
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  GGCAGCGGGCGGCGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCCGGTGAATTC  2457
```

```
RESULT 2
AX455915
LOCUS       AX455915                2463 bp    DNA             linear   PAT 06-JUL-2002
DEFINITION  Sequence 31 from Patent WO0204493.
ACCESSION   AX455915
VERSION     AX455915.1  GI:21714900
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE     Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 31 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
     source          1..2463
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PR975YM"
ORIGIN
```

```
Query Match        99.2%;  Score 2436.2;  DB 6;  Length 2463;
Best Local Similarity  99.6%;  Pred. No. 3.6e-246;
Matches 2454;  Conservative  0;  Mismatches  3;  Indels  6;  Gaps  1;

Qy     1  GTCGACGCGCACCATGGCCGAGGCCCATAGCCACCAGCGCCACCAGCCGCCAACATCCTGATGCAG  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GTGACGCGCACCATGGCCGAGGCCCATAGCCACCAGCGCCACCAGCCGCCAACATCCTGATGCAG  60

Qy    61  CGCAGCAACTTCAAGGGCCCCAGCGCCCCAGCAGCCCGCATCATCAAGTGCTTCAACTCGGGCAAGGAGGGC  120
           |||||||||||||(|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CGCAGCAACTTCAAGGGCCCCAGCGCCCCAGCAGCCCGCATCATCAAGTGCTTCAACTCGGGCAAGGAGGGGC  120

Qy   121  CACATCGCCCGCAACTGCCGCGCGCCCCCCCGCGCCCCCGCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CACATCGCCCGCAACTGCCGCGCGCCCCCCCGCGCCCCCGTACGTGGGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG  180
```

```
Qy   2395 TACGTGGGCAGCGGCGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2454
Db   2401 TACGTGGGCAGCGGCGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460

Qy   2455 TTC 2457
Db   2461 TTC 2463
```

RESULT 3
```
LOCUS        AX455914                 2469 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION   Sequence 30 from Patent WO0204493.
ACCESSION    AX455914
VERSION      AX455914.1  GI:21714899
KEYWORDS     .
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1
  AUTHORS    zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE      Polynucleotides encoding antigenic hiv type c polypeptides,
             polypeptides and uses thereof
  JOURNAL    Patent: WO 020493-A 30 17-JAN-2002; University of Stellenbosch (ZA)
             CHIRON CORPORATION (US) ;
FEATURES             Location/Qualifiers
     source          1..2469
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PR975(+)"

ORIGIN

Query Match             98.3%;  Score 2415.4;  DB 6;  Length 2469;
Best Local Similarity   99.3%;  Pred. No. 5.3e-244;
Matches 2451;  Conservative  0;  Mismatches  6;  Indels  12;  Gaps  2;
```

```
Qy     1 GTCGAGCGCCACCATGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GTCGAGCGCCACCATGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60

Qy    61 CGCAGCAACTTCAAGGGCCATCAAGTGCTTCAACTGCGGCAAGAGGGGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CGCAGCAACTTCAAGGGCCATCAAGTGCTTCAACTGCGGCAAGAGGGGC 120

Qy   121 CACATCCCCGCAACTGCCGCCCCGCGGCCCCCGCGGCCCCCGCAAGAAGGGCGTGCTGAAGTGCCGGCAAGGAG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACATCCCCGCAACTGCCGCCCCGCGGCCCCCGCGGCCCCCGCAAGAAGGGCGTGCTGAAGTGCCGGCAAGGAG 180

Qy   181 GGCCACCAGATGAAGGACTGCACCGCCGCCGCCAACTTCTTCCGCGAGGAGACCTGGCC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCCACCAGATGAAGGACTGCACCGCCGCCGCCAACTTCTTCCGCGAGGAGACCTGGCC 240

Qy   241 TTCCCCAGGGCAAGGCCAGGTCCCCGGCAACCCCGCCGAGGAGAACCGCGGCCAACAGCCCCACC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCCCAGGGCAAGGCCAGGTCCCCGGCAACCCCGCCGAGGAGAACCGCGGCCAACAGCCCCACC 300

Qy   301 AGCCGGCGAGCTGCAGTGCGCGGCGACAACCCCCGCAGCGAGGCGCCCCGCAGCGAGGCCGCCAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCCGGCGAGCTGCAGTGCGCGGCGACAACCCCCGCAGCGAGGCGCCCCGCAGCGAGGCCGCCAG 360

Qy   361 GGCACCCTGAACTTCCCCAGATCACCCGTGGCAGCGGCATCAAGGTG 420
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCACCCTGAACTTCCCCAGATCACCCGTGGCAGCGGCATCAAGGTG 420

Qy   421 GGCGGCCAGCTGAAGGAGCCCCGACGACCAAGGAGCCCGACGACCGTGCTGGAGGAG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGCCAGCTGAAGGAGCCCCGACGACCAAGGAGCCCGACGACCGTGCTGGAGGAG 480

Qy   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540
```

```
Db   1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGGCTGGGCTTCACCACCCCCGAC 1320

Qy   1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACAAG 1374
           ||||||||||||||||||||||||||            |||||||||||||||||||||
Db   1321 AAGAAGCACCAGAAGGAGCCCCCTTCCTGCCCGGCCGGCCTACGAGCTGCACCCCGACAAG 1380

Qy   1375 TGGACCGTGCAGCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG 1434
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG 1440

Qy   1435 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCCATCAAGGTGCGCCAG 1494
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCCATCAAGGTGCGCCAG 1500

Qy   1495 CCGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1554
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy   1555 GCCGAGTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1614
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1620

Qy   1615 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGACC 1674
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGACC 1680

Qy   1675 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACGTACGCCAAGATGCGC 1734
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACGTACGCCAAGATGCGC 1740

Qy   1735 ACCGGCCCACCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1794
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACCGGCCCACCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800

Qy   1795 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCTGCCCATCCAGAAGGAGACCTGG 1854
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCTGCCCATCCAGAAGGAGACCTGG 1860

Qy   1855 GAGAACCGGTGACGACTACTGGCAGGCCACTGCGAGGTCGGAGTTCGTGAAC 1914
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGAACCGGTGACGACTACTGGCAGGCCACTGCGAGGTCGGAGTTCGTGAAC 1920

Qy   1915 ACCCCCCCTGGTGGTACAGCTGGAAGCTCGAGAGCGAGGCCATCATCGGCGCCAG 1974
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACCCCCCCTGGTGGTACAGCTGGAAGCTCGAGAGCGAGGCCATCATCGGCGCCAG 1980

Qy   1975 ACCTTCTACGTGGACGGGCGCCGCAACCGCGAGAACTCGGCGAGGCCGGCTACGTG 2034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACCTTCTACGTGGACGGGCGCCGCAACCGCGAGAACTCGGCGAGGCCGGCTACGTG 2040

Qy   2035 ACCGACCGGGCCGGCAGAAGATCGTGACCGACCTGAGCCTGAACCACCAGAGAGACCGGAG 2094
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACCGGGCCGGCAGAAGATCGTGACCGACCTGAGCCTGAACCACCAGAGAGACCGGAG 2100

Qy   2095 CTGCAGGGCCATCCAGTCGGCCTGGCCAGCAGCGGCAGGTGAACATCGTGACCGAC 2154
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGGCCATCCAGTCGGCCTGGCCAGCAGCGGCAGGTGAACATCGTGACCGAC 2160

Qy   2155 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGAGCGAGCGAGCTGGTG 2214
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGAGCGAGCGAGCTGGTG 2220

Qy   2215 AACCAGATCATCGGCGCAGCAGTGATCAAGAGAGAAGTGTACCTGGCCTGGGTGCCCGCC 2274
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCAGTGATCAAGAGAGAAGTGTACCTGGCCTGGGTGCCCGCC 2280

Qy   2275 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTCGTGTGAGCAGGGCATCCGCAAG 2334
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTCGTGTGAGCAGGGCATCCGCAAG 2340

Qy   2335 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGAAGCCCAAGATGATCGGCGGCTTCATCAAG 2394
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGAAGCCCAAGATGATCGGCGGCTTCATCAAG 2400
```

```
Qy   1609 GTGTACTACGACCCCAGCAAGGAACTGGTGCCGAGATCCAGAGGACCAGGGCCACGACCAG 1668
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTGTACTACGACCCCAGCAAGGAACTGGTGCCGAGATCCAGAGGACCAGGGCCACGACCAG 1680

Qy   1669 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCCAAG 1728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCCAAG 1740

Qy   1729 ATGCGACCGCCACCAACGACGTGAAGAGCAGCTGACCGAGGCCGTGCCAGAAGATCGGC 1788
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATGCGACCGCCACCAACGACGTGAAGAGCAGCTGACCGAGGCCGTGCCAGAAGATCGGC 1800

Qy   1789 ATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAGGAG 1848
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAGGAG 1860

Qy   1849 ACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC 1908
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC 1920

Qy   1909 GTGAACACCCCCCGTGGTGAAGCTGGTGTTACCAGTGGAAGAGCCCATCATCGGC 1968
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GTGAACACCCCCCGTGGTGAAGCTGGTGTTACCAGTGGAAGAGCCCATCATCGGC 1980

Qy   1969 GCCGAGACCTTCTACGTGGACGCGGCCGCCAACCGGAGACCAAGATCGGCAAGGCCGGC 2028
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GCCGAGACCTTCTACGTGGACGCGGCCGCCAACCGGAGACCAAGATCGGCAAGGCCGGC 2040

Qy   2029 TACGTGACCGACGGGGCCGGCAGAGATCGTGAGCGCTGACCGAGACCAACCAGAAG 2088
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TACGTGACCGACGGGGCCGGCAGAGATCGTGAGCGCTGACCGAGACCAACCAGAAG 2100

Qy   2089 ACCGAGCTGCAGGCCATCCAGTCTGGCCCTGCAGCAGCGGCAGGTGAACATCGTG 2148
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ACCGAGCTGCAGGCCATCCAGTCTGGCCCTGCAGCAGCGGCAGGTGAACATCGTG 2160

Qy   2149 ACCGACAGCAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGAAGAGGAGAGCGAG 2208
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACCGACAGCAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGAAGAGGAGAGCGAG 2220

Qy   2209 CTGGTGAACCAGATCATCAAGAGCCAGCTGATCAAGAGAAGAGAAGGTGTTACCTGAGCTGGGTG 2268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CTGGTGAACCAGATCATCAAGAGCCAGCTGATCAAGAGAAGAGAAGGTGTTACCTGAGCTGGGTG 2280

Qy   2269 CCCGCCCAGAGGGCATCCGGCAGCGACCTGGAGATCGGC 2328
          |||||||||||||||||||||||||||||||||||||||
Db   2281 CCCGCCCAGAGGGCATCCGGCAGCGACCTGGAGATCGGC 2340

Qy   2329 CGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGAC 2388
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGAC 2400

Qy   2389 GACCTGTACGTGGCGCAGGCGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC 2448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GACCTGTACGTGGCGCAGGCGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC 2460

Qy   2449 GGTGAATTC 2457
          |||||||||
Db   2461 GGTGAATTC 2469
```

RESULT 4
BD263705
```
LOCUS       BD263705                2300 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Improved expression of HIV polypeptides and production of
            virus-like particles.
ACCESSION   BD263705
VERSION     BD263705.1  GI:33073473
KEYWORDS    JP 2002533124-A/72.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 2300)
```

```
Qy    541 GTGCGCCAGTACGACCAGATCCTGATCGAGCATGCGGGCAAGGAGGCCATCGCCACCGTG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGCGGCCAGTACGACCAGATCCTGATCGAGCATGCGGGCAAGGAGGCCATCGCCACCGTG 600

Qy    601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCGCAACATCGTGCCCAGCTGGCGGTGC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCGCAACATCGTGCCCAGCTGGCGGTGC 660

Qy    661 ACCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGCGGCATG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGCGGCATG 720

Qy    721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCC 780

Qy    781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840

Qy    841 AACACCCCGTGTTCGCCATCAAGAGAGGACAGCACCAAGCTGGCGCAAGCTGGTGGAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACACCCCGTGTTCGCCATCAAGAGAGGACAGCACCAAGCTGGCGCAAGCTGGTGGAC 900

Qy    901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC 960

Qy    961 CCCGCCGGCCTGAAGAAGAAGAAGAGCCGGCGTTGACCGTGCTGGGGCGACGGCCTACTTC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCCGCCGGCCTGAAGAAGAAGAAGAGCCGGCGTTGACCGTGCTGGGGCGACGGCCTACTTC 1020

Qy   1021 AGCGTGCCCCTGGACGAGGACTTCCGCACCTACGTACACCGGCTTCACCATCCCCAGCATCAAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGCGTGCCCCTGGACGAGGACTTCCGCACCTACGTACACCGGCTTCACCATCCCCAGCATCAAC 1080

Qy   1081 AACGAGACCCCCGGCATCCGGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACGAGACCCCCGGCATCCGGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140

Qy   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCGCAACCCC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCGCAACCCC 1200

Qy   1201 GAGATCGTGATCTACCA-----GGCTCCGTACGTGGGCAGCGACCTGGAGATCGGC 1254
          ||||||||||||||||||     |||||||||||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCAGTACAGGGCTCCGTACGTGGGCAGCGACCTGGAGATCGGC 1260

Qy   1255 CAGCACCCGGCGCCAAGATCGAGGAGCTGCGGCAAGCACCTGCTGCGGTGGCTTCACCACC 1314
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CAGCACCCGGCGCCAAGATCGAGGAGCTGCGGCAAGCACCTGCTGCGGTGGCTTCACCACC 1320

Qy   1315 CCCGACAAGAAGCACCAGAAGGAGCCCCCTTCT------CGAGCTGCACCCC 1368
          ||||||||||||||||||||||||||||||||      |||||||||||||
Db   1321 CCCGACAAGAAGCACCAGAAGGAGCCCCCTTCCCGTGGATGGGCTACCAGAGCTGCACCCC 1380

Qy   1369 GACAAGTGGACCGTGCAGCCTGCAGCTGCGGCCCATGCAGCTGCAGAGGAGACGGTGAACGAC 1428
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACAAGTGGACCGTGCAGCCTGCAGCTGCGGCCCATGCAGCTGCAGAGGAGACGGTGAACGAC 1440

Qy   1429 ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG 1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG 1500

Qy   1489 CGCCAGCTGTGTCAAGCTGCTGCGCGGCGCGCCCAAGGCCCTGACCGAGATCGTGCCCCTGACC 1548
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CGCCAGCTGTGTCAAGCTGCTGCGCGGCGCGCCCAAGGCCCTGACCGAGATCGTGCCCCTGACC 1560

Qy   1549 GAGGAGGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGCCCGAGATCCTGCCCGTGCACGGC 1608
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GAGGAGGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGCCCGAGATCCTGCCCGTGCACGGC 1620
```

```
AUTHORS   Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
          Greer,C., Selby,M. and Walker,C.
TITLE     Improved expression of HIV polypeptides and production of
          virus-like particles
JOURNAL   Patent: JP 2002533124-A 72 08-OCT-2002;
          CHIRON CORP

COMMENT
          OS    Artificial Sequence
          PN    JP 2002533124-A/72
          PD    08-OCT-2002
          PF    30-DEC-1999 JP 2000591193
          PR    31-DEC-1998 US   60/114495,01-DEC-1999 US   60/168471 PI
          SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
          KARIN HARTOG,
          PI HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
          C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
          C12N5/10,
          PC  C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/00 CC
          Description of Artificial Sequence: FS(-).protmod.RTopt.YMWM FH
          Key     Location/Qualifiers
          FT    source     1..2300
          FT            /organism='Artificial Sequence'.

FEATURES             Location/Qualifiers
     source          1..2300
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"

ORIGIN

Query Match          83.0%;   Score 2040;   DB 6;   Length 2300;
Best Local Similarity 93.6%;   Pred. No. 1e-204;
Matches 2153;   Conservative   0;   Mismatches 135;   Indels  12;   Gaps   2;
```

```
Qy    170 GCGGCAAGGAGGGCCACCAGATGAAGACTGACCGAGCTGCACCGAGCTGCCAGGCCAACTTCTTCCGCG  229
              |||  |||| || ||||  ||||||||  || ||  ||   ||    ||   || || ||| ||||||||
Db      1 GCGGCGCCGCAAGGAACACCCAAATGAAGAATGAAGATTGCACTGAGAAGACAGGCTAATTTCTTTCCGCG   60

Qy    230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGGGCCGCGAGAGTCGAGCAGCAGCCCGGCCA      289
              ||||||||||||||  |||||||||||||||||||||||||||||| || ||||| |||||||||
Db     61 AGGACCTGGCCTTCTGCGAGGGCAAGGCCCGCGAGGGCCGCGAGAGTCAGCAGCAGCCCGGCCA       120

Qy    290 ACAGCCCCACCAGCCGCGAGCTGCGAGGTGCGCGCG-----ACAACCCCGCCAGCGAGG         343
              ||||||| |||||||||||||||| ||||| |||||     ||||||||||||||||||||
Db    121 ACAGCCCCACCAGCCGCGAGCTGCCAGGTGTGCGGCGGCGAGAGACAACAGCCTGAGCGAGG       180

Qy    344 CCGGCGCCGCAGCGCGCCAGGCCACCCTG-----AACTTCCCCAGATCACCCTGTGGCAGC       397
              ||||| ||||||||||||||||||||||     |||||||||||||||||||||||||||
Db    181 CCGGCGCCGACGCCGCGCCAGGGCACCCGTGAGCTTCAACTTCCCCAGATCACCCTGTGGCAGC    240

Qy    398 GCCCCTGTGAGCATCAAGGTGGCGGCAGATCGGCGGCCGGCTCAAGGATGATCG          457
              |||||||||||||||||||||||||||| |||||| ||||||| ||||||||||
Db    241 GCCCCTGGTGACCATCAAGGTGCGGGGCCAGATCGGCGGCCAGCTGAAGGATGATCG       300

Qy    458 CCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCCAAGATGATCG       517
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db    301 CCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCCAAGATGATCG    360

Qy    518 GCGGCATCGGCGGCGGCTTCATCAAGGTCGCGCACCGTGCTGCGCGCCAGTGCGCGGGATCGGCTGCG      577
              |||||||||||||||||||||||||||| ||||| ||||||||||||| |||||||||| || |||
Db    361 GCGGGATCGGCGGCTTCATCAAGGTGCGCGCACCGTGCTGCGCGCCAGTGCGCGGGGATCTGCG    420

Qy    578 GCAAGAGGCCATCGGCACCGTGCTGACCCTGCCGACATCATCGGCCCA      637
              ||||| ||||||||||||||||||| |||| ||||| |||||||||||
Db    421 GCCACAGGCCATCGGCACCGTGCTGCTGCTCCGCGTGAACATCATCGGCGGCCA       480

Qy    638 ACATGCTGACCCAGCTGGCGTCGACCCTGCCGCACCCCATCAGCCCATCAGCGACACCGTGC      697
              |||| ||| ||||| ||||||||||| ||||||||||||||||| ||||||||||||||||
Db    481 ACCTGCTGACCCAGCTGGCGTCGACCCTGCCGCACCCCATCAGCTTCCCCATCAGCGACACCGTGC    540

Qy    698 CCGTGAAGCTGAAGCCCGGCACGGCGACGGGCGATGACGAGCGGGTGAAGCAGCAGTGAAGCAGTGCGCTGACCGAGG      757
              |||||||||||||||||||||||||| |||| |||| |||||||||||||||| |||||||||
Db    541 CCGTGAAGCTGAAGCCCGGGGATGGGACGGGGCGGGCGGCCCGGACGGGCGACCCCTGACCGAGG    600

Qy    758 AGAAGATCAGGGCCCTGACCGCCATCGCCGAGGAGATGGAGAAGGAGGGGAAGATCACCA      817
              |||||| || || || || |||| || ||||||||||| |||||||| ||||||||||
Db      ...
```

```
Qy    601 AGAAGATCAGGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGATGGAGAAGATCAGCA      660
          |||||||||||||||||  ||||| || |||||||||||||||||   |||||||||||
Db    818 AGATCGGCGCCGGAGAACCCTACACACCCCGTGTTCGCCATCACAGAAGAAGGACAGCA      877

Qy    661 AGATCGGCGCCGGAGAACCCTACACACCCCGTGTTCGCCATCACAGAAGAAGGACAGCA      720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      ...

Qy    721 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG      780
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    878 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG      937

Qy    781 AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC      840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    938 AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC      997

Qy    841 TGGACGTGGGCGACGCCTACTTCAGGGTGCCCCTGGACGAGGACTTCCGCACCTACACCG      900
          |||||||||||| |||||||||||| ||||||||||||||||||||| |||||||||||
Db    998 TGGACGTGGGCGACGCCTACTTCAGGGTGCCCCTGGACGAGGACTTCCGCACCTACACCG     1057

Qy    901 CCTTCACCATCCCCAGCATCAACAACGAGACCCCGGCATCCGCTACCAGTACACACGTGC      960
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   1058 CCTTCACCATCCCCAGCATCAACAACGAGACCCCGGCATCCGCTACCAGTACACACGTGC     1117

Qy    961 TGCCCCAGGGCTGGAAGGGCAGCCCGCCCATCTTCCAGAGCATCATGACCAAGATCCTGG     1020
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   1118 TGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCCAGAGCATGATGACCAAGATCCTGG     1177

Qy   1021 AGCCCTTCCGCGCCCGAAGCAGAACCCGACATCGTGATCGTGATCGTGACGCACCTGCTGC     1080
          ||||||||||||| |||||||| ||||||| |||||||||||||||  ||||||||||||
Db   1178 AGCCCTTCCGCGCCCAAGCAGAACCCGACAGATCGTGATCGTGGGCA      1237

Qy   1081 GCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGC     1140
          ||||| ||||||||||| ||||||||||||||||||||||||| ||||||||||||||
Db   1238 GCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGC     1297

Qy   1141 GCTGGGGCTTCACCACCCCAGACAAGAAGCACCAGAAGGAGCCCCCATCCTTCCTGCCCATCG     1200
          |||| ||||||||| ||  ||||||  ||  || |||||||| ||||||||||||||||
Db   1298 GCTGGGGCTTCACCACGCCCGACAAGAAGCACCAGAAGGAGCCCCCATCCTTCCTGCCCATCG     1357

Qy   1201 AGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGA     1260
          |||||||||||||||||||| ||||| ||||| ||||||||||||||||||||||||||
Db   1358 AGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGA     1417

Qy   1261 CCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCCAGATCTACCCCG     1320
          |||||||| ||||||||||||||||||||||||||| |||||||| |||||||||||
Db   1418 CCGTGAACGACATCAAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCG     1477

Qy   1321 GCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGAGGTGA     1380
          ||||  ||||| ||||||||||||||| |||||||||||||| |||||| ||| |||||
Db   1478 GCATCAAGGTGCGCGAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCAGCGAGC     1537

Qy   1381 TCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGC     1440
          || || ||||||||||| ||||||||||||||| |||||||| ||||| |||| |||||
Db   1538 TGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGC     1597

Qy   1441 CGGTGCACGGGGTGTACTACGACCCCAGCAAGGACCTGGTCGCCGAGATCCAGAAGCAGG     1500
          ||||||||| ||||||||||||||||| |||||||| ||||||||||||||||||||||
Db   1598 CGGTGCACGACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGG     1657

Qy   1501 GCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCA     1560
          |||| ||||||||| || |||||||| ||||||||||||| |||| || ||||||| ||
Db   1658 GCCAGGACCTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCA     1717

Qy   1561 AGTACGCGCGCATGAGGGGCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGC     1620
          || ||||| || ||||||| || ||||| |||| ||||||||||||||||||||| ||
Db   1718 AGTACGCCAAGATGCGCGGCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGC     1777

Qy   1621 AGAAGGTGAGCACCGAGAGCATCGTGATCTGGGGCAAGATCCCCAAGTTCAAGCTGCCCA     1680
          |||| ||||| ||||||||||||| || ||||| |||||||| ||||| ||||||||||
Db   1778 AGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGGCTGCCCA     1837

Qy   1838 TCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCG     1897
          ||||| ||||||| |||||||||||| ||||| |||||||||||||||||||||||||||
Db      ...
```

```
Db   61 AGGACCTTGGCCTTCCTGCAGGCGAAGGCCGCCGGAGTTCAGCAGCAGCAGACCACCGCCA  120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  120 TCCAGGAAGGAGGACCTTGGCGAGGAGGCCTGGTGGAGTACTGGCAGGCCACCTGGATCCCCG 1740

Db  290 ACAGCCCACCAGCCGCGAGCTGCAGGTGCCGCGCG------ACAACCCCGAGCGGAGG      343
        ||||||||||||||||||||||||||||||||||     |||||||||||||||||||
Qy 1898 AGTGGGAGTTCGTGAACACCAGCCCCCCGGTGAAGCTGTGGTGAAGCTGTTCAGCGAGGAAGGAGC 1957

Db  121 ACAGCCCACCCGCGCGGCGAGCTGCAGGTGTGGGGGCGGAGAACAACAGCCTGAGCGGAGG  180
        |||||||||||||||||||||||||||||
Qy 1741 AGTGGGAGTTCGTGAACACCAGCCCCCCCGGTGAAGCTGTGGTGAAGCTGTTCAGCGAGGAAGGAGC 1800

Db  344 CCGGCGCGAGCGCCAGGGCACCTG-----AACTTCCCCAGATCACCCTGTGCAGC      397
Qy 1958 CCATCATCGGCGCCGAGACCTTCTACGTGACGGCGCGCCAACCGCGAGACCAAGAGATCG 2017

Db  181 CCGGCGCCGAGCGCCAGGGCACCTTCAACTTCCCCCAGATCACCCTGTGCTGGCAGC  240
Qy 1801 CCATCGTGGGCGCCGAGACCTTCTACGTGACGGCGCGCCAACCGCGAGACCAAGCTGG 1860

Db  398 GCCCCTGGTGAGCATCAAGGTGGGGGCCGGCAGAGATCGTGAGCCTGACCGAGACCA  457
Qy 2018 GCAGGCGCCGGCTACGTGACGACCGGGGCCGGGCCGCAGAAGATCGTGAGCCTGACCGAGACCA 2077

Db  241 GCCCCCTGGTGACCATCAGGATCGGCGGCGGCGGCTCAAGGAGGCGGCGGTGCTCGACACCACCGCG  300
Qy 1861 GCAAGGGCCGGCTACGTGACGACCGGGGCCGGGCCGCAGAAGGTCGTGAGCATCGCCGACCACCA 1920

Db  458 CCGAGACCACCGTGCTGGAGGAGGTGAGCCTGCCCGGCCAGAGTGAGCCTGACGGCCCAAGATGATCG  517
Qy 2078 CCAACCAGGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCCGCAGCGAGG 2137

Db  301 CCGACGACACCGTGCTGCAGGCGAGATGAAGCTGGAGGAGCAGCAGCCCAAGATGATCG  360
Qy 1921 CCAACCAGGAAGACCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACAGCCGCAGGCCGCTGGAGG 1980

Db  518 GGGCATCGGCGGCTTCATCAAGGTGCGCCCAGTACGACCCAGATCCTGATCGAGATCTCGG  577
Qy 2138 TGAACATCGTGACCGACAGCAGCCAGTACGCCCTGGCATCATCCAGGCCCGACAAGA 2197

Db  361 GCGGGATCGGGGCTTCATCAAGGTGCGCCCAGTACGACCCCAGATCCTGGAGATCTCGC  420
Qy 1981 TGAACATCGTGACCGACAGCAGCCAGTACGCCCTGGGCATCATCCAGCCCCGACAAGA 2040

Db  578 GCAGAGAGGCCATCGGCACCGTGCTGATCGGCCCGAACATCATCGGCCCGCCA  637
Qy 2198 GCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACC 2257

Db  421 GCCACAGGCACCGTGCTGGTGGGCCCACCCCCGTGAACATCATCGGCCCATCATCGGCCCGCCA  480
Qy 2041 GCCGAGAGCGAGCTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACC 2100

Db  638 ACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGC  697
Qy 2258 TGAGCTGGGTGGCCGCCCACAAGGCCATCGGGGCCCACAAGGCGCAAGCTGGTGA 2317

Db  481 ACCTGCTGACCCAGCTGTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGA  540
Qy 2101 TGGCCCTGGGCTGCCCGCCACAGAGGGCCATCGGCGGGCATCGGGGCCCACAAGGCCGCAAGCTGGTGA 2160

Db  698 CCGTGAAGCTGAAGCCCCAAGGTGACGACGGCCCCAAGGTCAGTGCCCTGACCGAG  757
Qy 2318 GCAAGGGCATCCGCAAGGTGCTGTTCCTGACCAGGCGGCCATCGATGGGCGGCATCGTGATCTACC 2377

Db  541 CCGTGAAGCTGAAGCCCCAAGGTCAGTGCCCTGACCGAG  600
Qy 2161 GCGCGCGGCCATCCGCAAGGTGCTGTTCCTGAACGGCATCGATGGGCGGCATCGTGATCTACC 2220

Db  758 AGAAGATCAAGGCCCTGACCGCCATCTGCGACGAGATCTGCACCACCCCGTGTTCGCCATCA  817
Qy 2378 AGTACATGGACACACCTGTACGTGGGCAGCGGCGGCGGCGGGCCCTAGGATCGATTAAAAAGCTTCCCG 2437

Db  601 AGAAGATCAAGGCCCTGACCGCCATCTGCGACGAGATCTGCACCACCCCGTGTTCGCCATCA  660
Qy 2221 AGTACATGGACACACCTGTACGTGGGCAGCGGCGGCGGGCCCTAGGATCGATTAAAAAGCTTCCCG 2280

Db  818 AGATCGGCCCCGAGAACACCCTACACACCCCGTGTTCGCCATCAAGAGAAGGACAGCA  877
Qy 2438 GGGCTAGCACCGGTGAATTC 2457

Db  661 AGATCGGCCCCCGAGAACCCCTACACAACACCCCGTGTTCGCCATCAAGAGAAGGACAGCA  720
Qy 2281 GGGCTAGCACCGGTGAATTC 2300

Db  878 CCAAGTGGCGCAAGCTGGTGGACTTCCGGCGACTTCGCCACCCGACCACCTTCTGGG  937
Db  721 CCAAGTGGCGCAAGCTGGTGGACTTCCGGCGACTTCGCCACCCGACCACTTCTGGG  780

Db  938 AGGTGCAGCTGGGCATCCCCCGGCCTGAAAGAAGAGAAGCGTGACCGGTGC  997
Db  781 AGGTGCAGCTGGGCATCCCCCCACCCCCGGCCTGAAAGAAGAGAAGCGTGACCGGTGC  840

Qy  998 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGGACGAGAGACTTCCGCAAGTACACCG 1057
Db  841 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGGACGAGGACTTCCGCAAGTACACCG  900

Qy 1058 CCTTCACCATCCCCAGCATCAACACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 1117
Db  901 CCTTCACCATCCCCAGCATCAACACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC  960

Qy 1118 TGCCCCAGGGGCTGGAAGGGCAGCCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
Db  961 TGCCCCAGGGGCTGGAAGGGCAGCGCCCCCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020

Qy 1178 AGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCTGTACGTGGGCA 1237
Db 1021 AGCCCTTCCGCAAGCAGACAGAACCCCCGACATCGTGATCTACCAGGCCCCCTGTACGTGGGCA 1080

Qy 1238 GCGACCTGGAGATCGGCGCCAAGATGCAGGCTGCGCAAGCACCTGCTGC 1297
Db 1081 GCGACCTGGAGATCGGCGCCACCACCCAGCACCCAAGATCGCGCGGCACCACCTGCTGC 1140

Qy 1298 GCTGGGCTTCACCACCCCGGACAGAAGAGACCAGGAAGGAGCCCCCTTCCTGCCCATCG 1357
```

Query Match   83.0%; Score 2040; DB 6; Length 2300;
Best Local Similarity 93.6%; Pred. No. 1e-204;
Matches 2153; Conservative 0; Mismatches 135; Indels 12; Gaps 2;

```
Qy  170 GCGGCAAGGAGGAGCCCACCAGCATGAAGGACTGCACCGAGCGCCAGGCCACTTCTTCCGG  229
Db    1 GCGGCCGCGAAGGAGACCACCAAATGAAAGATTGCACTGAGACAGACCAGGCTARTTCTTCCCCG   60
Qy  230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCGCCGGAGTTCCCAGCAGCAGAGAACCGCGGCCA  289
```

```
Qy    2438  GGGCTAGCACCGGTGAATTC  2457
            ||||||||||||||||||||
Db    2281  GGGCTAGCACCGGTGAATTC  2300

RESULT 6
LOCUS       BD263704              2306 bp    DNA      linear   PAT 17-JUL-2003
DEFINITION  Improved expression of HIV polypeptides and production of
            virus-like particles.
ACCESSION   BD263704
VERSION     BD263704.1  GI:33073472
KEYWORDS    JP 2002533124-A/71.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 2306)
  AUTHORS   Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
            Greer,C., Selby,M. and Walker,C.
  TITLE     Improved expression of HIV polypeptides and production of
            virus-like particles
  JOURNAL   Patent: JP 2002533124-A 71 08-OCT-2002;
            CHIRON CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002533124-A/71
            PD   08-OCT-2002
            PF   30-DEC-1999 JP 2000591193
            PR   31-DEC-1998 US  60/114495,01-DEC-1999 US   60/168471 PI
            SUSAN BARNETT,JAN ZUR MEGEDE,INDRSH SRIVASTAVA,YING LIAN, PI
            KARIN HARTOG,
            PI   HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
            C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
            C12N5/10,
            PC   C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
            Description of Artificial Sequence: FS(-).protmod.RTopt.YM FH  Key
            Location/Qualifiers
            1..2306
FT          source
FT                    /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1..2306
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match        82.2%;  Score 2019.2;  DB 6;  Length 2306;
Best Local Similarity   93.2%;  Pred. No. 1.5e-202;
Matches 2150;  Conservative   0;  Mismatches 138;  Indels   18;  Gaps    3;

Qy     170  GCGGCAAGGAGGAGGCCACCAGATGAAGGACTGACCAGCGCCAGGCCAACTTCTTCCGCG  229
            |||||||||||||||| ||||||||||||||||| |||||||||| |||||| |||||||
Db       1  GCGGCGCCGGAAGGAGCACCAAATGAAGAGATTGCACTGCACGAGGACCAGGCGAGC    60

Qy     230  AGGACCTGCCTTCCCCCAGGGCACATCGCCCTGGGCATCATCCCCGAGGCAGACCGGCCA  289
            |||||||||||||||||||||||| ||||| ||||||||||||||||||||||||||||||
Db      61  AGGACCTGCCTTCCTCGAGGCAGGGCCAAGGCCGCCGCAAGGCCAAGCCAGCAGACCGGCGCCA   120

Qy     290  ACAGCCCACCACCAGCCGCCAGCTGCAGGTGCGCGGCG-----ACAACCCCGCAGGCGAGG  343
            |||||||||||||||||||||||||||| |||||||||     ||||||||||||||||||
Db     121  ACAGCCCACCACCCGCCGCCAGCTCGCAGGTGTGGGCGGCGGAGAACAACAGCAGCCTGAGCGAGG   180

Qy     344  CCGGCGGCCAGCCGCCGCCAGGCACCCTG------AACTTCCCCAGATCACCCTGTGGCAGC  397
            |||||||||||||||||||||||||||||      |||||||||||||||||||||||||||
Db     181  CCGGCGCCGACCGCCGCCAGGGCACCCTCAACTTCCCCGAGCTTCAACTTCCCCAGATCACCCTGTGGCAGC   240

Qy     398  GCCCCTGGTGAGCATCAAGGTGGGCGGCATCATGCGGCGGCATCATGCGGCATCGTGATCTACC  457
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  GCCCCTGGTGACCATCAGGATCGGCGGCATCATCGTGATCTACC   300

Qy     458  CCGACGACACCGGTGCTGGAGGAGTGAGAGCGGCCAAGTGGAAGCCCAAGATGATCG  517
            ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db     301  CCGACGACACCGTGCTGGGAGGGAGATGAACCTGGCGGCGGGCCCAAGTGGAAGCCCAAGATGATCG   360
```

```
Db    1141  GCTGGGGCTTCACCACCCCGACAAGAGAGCACCAGAGAAGGAGCCCCTTCTGCCCATCG  1200

Qy    1358  AGTGCACCCGACCAAGTGGACCGTGCCAGCTGCAGCCCATCGAGCTGCCCGAGGAGAGCTGGA  1417
            ||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db    1201  AGTGCACCCGACCAAGTGGACCGTGCCGCAGCCCATCATCGTGCCCGAGAAGGACAGCTGGA  1260

Qy    1418  CCGTGAACGACATCATCCAGAGAGTCGTGGCAAGCTGAACTGACTGGGCCAGCCAGATCTACCCCG  1477
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    1261  CCGTGAACGACATCATCCAGAGAGTGGGCCAGCCACCAAGCCCTGACCAGCCAGATCTACGCCG  1320

Qy    1478  GCATCAAGGTGCGCCAGTCTGCAAGCTCTGCCGGCGGCCAAGCCCTGACCGACATCG  1537
            |||||||||| |||||||||||||||| |||||||||||||||||||||||||
Db    1321  GCATCAAGGTGAAGCAGGTGCTGCAAGCTGTGCAAGCAGGCCCTGACCTGACCGACGTGA  1380

Qy    1538  TGCCCCTGACCGAGGAGCCGAGCTGGAGCTGGCTGGAGCTGGCTGCGAGATCCTGCGCAGC  1597
            |||||||||| ||||| |||||||||||||||||||||||||||||||||||||
Db    1381  TCCCCCTGACCGAGGAGCCGAGCTGGAGCTGGCTGGCGGAGAACCGCGAGATCCTGAAGGAGC  1440

Qy    1598  CCGTGCACGGCGGTGTACTACGACCCCAGCAAGGACCTGTGTGGCGAGATCCAGAAGCAGG  1657
            |||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db    1441  CCGTGCACGAGGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGG  1500

Qy    1658  GCCACGACCAGTGACCTACCAGATCTTACCAGAGACCCTTCAAGAACCTGAAGACCGGCA  1717
            |||||||||| ||||| ||||||||||||| ||||||| |||||||||||||||||||
Db    1501  GCCAGGCCCAGTGACCTACCAGATCTTACCAGAGAGCCCTTCAAGAACCTGAAGACCGGCA  1560

Qy    1718  AGTACGCCAAGATGCGCACCGCCACCACAACGACGTGAAGCAGCTGACCGAGGCCGTGC  1777
            |||||||||| |||||||||||||||||||| |||||||||||||||||||||||||||
Db    1561  AGTACGCCAAGATGCGCGGCGGCCATGCGCGGCCCACACCAAGACGTGAAGCAGCTGACCGAGGGCCGTGC  1620

Qy    1778  AGAAGATCGCCATGAGAGGCATCGGTGATCGTGGGGGCAAGACCCCCAAGTTCGCCTGCCCA  1837
            ||||| |||||||||| |||||||||||||||| |||||||||| ||||||||||||||||
Db    1621  AGAAGGTGAGCACCGAGAGCATCGTCGTGATCGTGGGCGCAAGATCCCCAAGTTCAAGTGCCCA  1680

Qy    1838  TCCAGAAGGACACCGTGGGACGACCTGGTGACCGACTACTGGCAGACTACTGGGATCCCCG  1897
            |||||||||||||| |||||  |||||||||||||||||||||  |||||||||||||||
Db    1681  TCCAGAAGGACACCCTGGAGGCCTGGTGGTGATGGAGTACTGCAGGGCCACCTGGGATCCCCG  1740

Qy    1898  AGTGGGAGTTCGTGAACACACCCCCCCTGTGAAGCTGTGGTACCAGCTGGAGAAGGAGC  1957
            ||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    1741  AGTGGGAGTTCGTGAACACACCCCCCCTGTGTGAAGCTGTGGTACCAGCTGGAGAAGGAGC  1800

Qy    1958  CCATCATCGGCGCCGAGACCTTCTACGTGGACGGCCGCGCCAACCGGCGAGACCAAGATCG  2017
            |||||||||| ||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    1801  CCATCGTGGGCGCCGAGACCTTCTACGTGGACGGCCGACGGCCGCGCCAACCGGCGAGACCAAGCTGG  1860

Qy    2018  GCAAGGCCGGCTACGTGACCGACCGGGCGCCGGCAGAGATCGTGAGCTGCAGCCGCAGACCA  2077
            |||||||||||||| ||||| |||||| |||||||||| ||| |||||| |||||||||||
Db    1861  GCAAGGCCGGCTACGTGACCGACCGACCGGGCCGGCAGAGGTGGTGAGCTGCTGGAGCATCGCCGCCACCA  1920

Qy    2078  CCAACCAGAAGAGACCGAGCTGCAGCCGCCATCCAGCTGGCCCTGACCGGCCGCAGCCGCAGG  2137
            ||||||||||||||||||||| |||||||| |||||||| ||||| |||||||||
Db    1921  CCAACCAGAAGAGACCGAGCTGCAGCTGCCATCCACCTGCCCTGCAGCAGCAGCGGCCTGGAGG  1980

Qy    2138  TGAACATCGTGACCGACAGCAGCCAGTACGCCCTGGGCATCATCCAGGCCCTGACCACAAGA  2197
            |||||||||||| |||||||||||| |||| |||||||||||||||||||||||||||||
Db    1981  TGAACATCGTGACCGACCAGCAGCCAGTACGCCCTGGGCATCATCCAGGCCCCGACACAAGA  2040

Qy    2198  GCGAGAGCGAGCTGTGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACC  2257
            |||||| ||||||||| |||||||| |||||||||||||||| ||||||||||||||||||
Db    2041  GCGAGAGCGAGCTGTGTGAGCAGCAGCTGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACC  2100

Qy    2258  TGAGCTGGGTGCCCGCCACAAGGCATCGGCAGATCGGCGCAACGAGCAGATCGACAAGCTGGTGA  2317
            ||||||||| ||||| |||||||||||||||||||||||| ||||||||||||| |||||||
Db    2101  TGGGCTGGGTGCCCCACGCCCCACCAAGGGCATCGGCGCCAAGAGGCAGCGTGACAAGCTGGTGA  2160

Qy    2318  GCAAGGGGCATCGCCAAGGTGCTGTTCCTGGAGCATCGGCGGCATCTGATCTACC  2377
            |||||| |||| ||||||||||||||||||| |||||||||| |||||||||||
Db    2161  GCGGCCGGCATCCGCCAAGGTGCTGTTCCTGAACCGGCATCGGCGGCATCTGATCTACC  2220

Qy    2378  AGTACATGGACGACCTGTACGTGGGCAGCGACCTAGATTGATTAAAAGCTTCCCG  2437
            |||||||| |||||||||||||||||||| |||||| ||||||||||||||||
Db    2221  AGTACATGGACGACCTGTACGTGGGCAGCGACCTAGATTCGATTGATTGATTAAAAGCTTCTTCCCG  2280
```

```
Qy  518  GCGGCATCGGCGCGGCTTCATCAAGGTAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGG  577
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GCGGGATCGGCGGGGCTTCATCAAGGTGCGGCAGTACGACCAGATCCCGTGAGATCTGCGG  420

Qy  578  GCAAGAGAGGCCATCGGCACCGTGCTGATCGGCGTCGTGGTGGCCCACCCGTGAACATCATCGGCGCCGCA  637
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GCCACAAGGCCATCGGCACCGTGGGTGCACCCGTGGGCCCCACCCGTGAACATCATCGGCGCCGCA  480

Qy  638  ACATGCTGACCCAGTCGGCTGGGCTGCACCTGAACTTCCCATCAGCCCATCGAGACCGTGC  697
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ACCTGCTGACCCAGATCGGCTGGGCTGCACCTGAACTTCCCATCAGCCCATCGAGACCGTGC  540

Qy  698  CCGTGAAGCTGAAGCCCGGCATGACGGCCCCAAGGTGAAGCAGTCGCCCCTGACCGAGG  757
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  CCGTGAAGCTGAAGCCCGGGATGGACGGCGCCCCAAGGTCAAGCAGTCGGCCCCTGACCGAGG  600

Qy  758  AGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGGCAAGGGCAAGATCACCA  817
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  AGAAGATCAAGGCCCTGGTGAGATCTGCACCGAGGAGATGGAGGAGGAGGGGCAAGATCAGCA  660

Qy  818  CCAAGTGGCGCAAGCTGTGACTTCCGTGTTCGGCATCGAACAAGGCCACCCCAGGGCACTTCTGGG  877
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  AGAAGATCAAGGCCCTGGTGAGATGGGAGTTCGTGAACACCCCGTGGTGAAGCTGTGGTACCAGCTGGAGA  720

Qy  878  CCAAGTGGCGCAAGCTGTGACTTCCGTGTTCGGCATCGAACAAGGCCACCCCAGGGCACTTCTGGG  937
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  CCAAGTGGCGCAAGCTGTGACTTCCGTGTTCGGCATCGAACAAGGCCACCCCAGGGCACTTCTGGG  780

Qy  938  AGGTGCAGCTGGGCATCCCCACCCCCCGCGCGGCCTGCCCTGACCGTGACCGTGC  997
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  AGGTGCAGCTGGGCATCCCCACCCCCCGCGCGGCCTGCCCTGACCGTGACCGTGC  840

Qy  998  TGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGAGTACGTGCAAGTACACCG  1057
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  TGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGAGTACGTGCAAGTACACCG  900

Qy  1058 CCTTCACCATCCCAGCATCAACGAGAACCCCGGCATCCGTACCAGTACAACGTGC  1117
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  CCTTCACCATCCCAGCATCAACGAGAACCCCGGCATCCGTACCAGTACAACGTGC  960

Qy  1118 TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGCAGCAGCAGCTCCTGG  1177
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGCAGCAGCAGCTCCTGG  1020

Qy  1178 AGCCCTTCCGCCCGAGATCGTGATCTACCAGGCCCCCTGTACGTGGGCA  1237
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AGCCCTTCCGCCCGAGATCGTGATCTACCAGGCCCCCTGTACGTGGGCA  1080

Qy  1238 GCGACCTGAGATCGGCCAGCACCGGCGCCAAGATCGGCGAGCACCTGCTGC  1297
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GCGACCTGAGATCGGCCAGCACCGGCGCCAAGATCGGCGAGCACCTGCTGC  1140

Qy  1298 GCTGGGGGCTTCACCACCCCGACAAGAGCACCACCCCGACAAGAGCACTCCTGCCCAT--  1355
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GCTGGGGCTTCACCACCCCGACAAGAGCACCACCCCGACAAGAGCACCTTCCTGTGGATGG  1200

Qy  1356 ----CGAGCTGCACCCCGACCAAGTGGCACCGTGCGCAGCTGGGACGGCCGCACCTTCCCGGCAGCTGGGCA  1411
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GCTACGAGCTGCACCCCGACCAAGTGGCACCGTGCGCAGCTGGGACGGCCGCACCTTCCCCGAGAAGGACA  1260

Qy  1412 GCTGGACCGTGAACGACATCCCCAGATCCAGAGAAGCTGGGCAAGCTGGAACTGGGCCAGCCAGATCT  1471
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GCTGGACCGTGAACGACATCATCCAGATCCAGAGAAGCTGGGCAAGCTGGGGCGCCCAGATCT  1320

Qy  1472 ACCCCCGGCATCGAGAAGGTGCGCCAGTCGTGCAAGGTCTCGCCGGCGGCCCTGACCG  1531
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 ACGCCCGGCATCAAGGTGAAGCAGGTTCGCGAGCTGGAGCTGGAGCTGGAGAGCCCTGACCG  1380

Qy  1532 ACATGCTGCCCCTGACCGACAAGCTGCCCGAGCTGGAGGCCCGGAGACCGCTGAGATCCTGC  1591
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 AGGTGATCCCCCTGACCGACAAGCTGCCCGAGCTGGAGGCCCGGAGACCGCTGAGATCCTGA  1440
```

```
Qy  1592 GCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGCAAGGACTCGTGTGGCCGGAATCCAGA  1651
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 AGGAGCCCGTGCACGAGTGTACTACGACCCCAGCAAGCAAGGACTCGGTGGCCGGAATCCAGA  1500

Qy  1652 AGCAGGGCCACCAGCCAGTGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGA  1711
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 AGCAGGGCCAGGGCCAGTGACCTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGA  1560

Qy  1712 CCGGCAAGTACGCCAAGATGGCGACCCCACCCCAACCCACCACCACACGAGCAGCGTGAACCGAGG  1771
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 CCGGCAAGTACGCCCGCATGGCCCGCATGGCGCGGCCGCCCACAGCAGCGTGAAGCAGTGGCCGAGG  1620

Qy  1772 CCGTGCAGAAGATGATCGCCATGGGCGCATCGTGATCTGGGCGACAAGTTCCGC  1831
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 CCGTGCAGAAGGTGAGCACCGAGCACCGGAGCATCGTGATCTGGGGCAAGTTCAAGC  1680

Qy  1832 TGCCCATCCAGAAGGAGACCTGGGAGACGTGGATCTGGACCGACTACTGGGCCACCTGA  1891
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGCCCATCCAGAAGGAGACCTGGGAGACGTGGATGTGGACTACTGGCAGGCCACCTGA  1740

Qy  1892 TCCCCGAGTTGGAGTTCGTGAACAACCCCTGGTGGACGGCGCCGCCCCGAGACCA  1951
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 TCCCCGAGTTGGAGTTCGTGAACAACCCCTGGTGGACGGCGCCGCCCCGAGACCA  1800

Qy  1952 AGGAGCCCATCATCGGCGCCGAGACTTCACGTGACCGGGCGCGAGAAGACCGCGAGACCGCGAGACCA  2011
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGGAGCCCATCATCGTGGGCCGCTACGTGACCGACGCCGGGGCGGCCGAGAAGTGGTGAGCCTGACCG  1860

Qy  2012 AGATCGGCAAGGCCGGCTACGTGACCGACGCCGGGGCGGCCGAGAAGATCGTGGTGAGCGCGCG  2071
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 AGCTGGGCAAGGCCGGCTACGTGACCGACGCCGGGGCGGCCGAGAAGTGGTGAGCAGCGCGGCA  1920

Qy  2072 AGACCACCAACAGGAAGAAGACCGAGCTGCAGGCCATCAGCTGCCCTGGCCCAGCAGAGCGGCGGCA  2131
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ACACCACCAACAGGAAGAACCGAGCTGCAGGCCATCAGCTGCCCTGGCCCAGCAGACCA  1980

Qy  2132 GCGAGGTGAACATCGTGACCGACCAGCTGCAGAACGTGACCGACCAGCTGCACGGCCCCAGCCCG  2191
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 TGGAGGTGAACATCGTGACCGACCAGCTGCACGCCCAGCTGCACGCCCCAGCCCG  2040

Qy  2192 ACAAGAGCGAGAGCCAGCTGGTGAACCAGATCATCGATCAAGAGAGGAGAAGG  2251
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ACAAGAGCGAGAGCCAGCTGGTGAGCCAGATCATCGATCAAGAGAGGAGAAGG  2100

Qy  2252 TGTACCTGAGCTGGGTGCCCCCACAAGGGCGCATCGGCGGCAACGAGCGATCGACAAGC  2311
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 TGTACCTGAGCTGGGTGCCCCCACAAGGGCGCATCGGCGGCAACGAGCGATCGACAAGC  2160

Qy  2312 TGGTGAGCAAGGGCCATCCGCAAGGTGCGTTCCTGGACGTGGCCGGCATCCTGA  2371
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TGGTGAGCAAGGGCCATCCGCAAGGTGCGTTCCTGGAACGCGGCCGGCATCCTGA  2220

Qy  2372 TCTACCAGTACATGACGACCTGTACGTGGGCACGCGGCGGCCCTAGGATCGATTAAAAGC  2431
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 TCTACCAGTACATGACGACCTGTACGTGGGCACGCGGCGGCCCTAGGATCGATTAAAAGC  2280

Qy  2432 TTCCCGGGGGCTAGCACCGGTGAATTC  2457
         |||||||||||||||||||||||
Db  2281 TTCCCGGGGGCTAGCACCGGTGAATTC  2306
```

```
RESULT 7
AR373387
LOCUS       AR373387          2306 bp    DNA     linear    PAT 18-DEC-2003
DEFINITION  Sequence 82 from patent US 6602705.
ACCESSION   AR373387
VERSION     AR373387.1  GI:40075490
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2306)
```

```
TITLE        Expression of HIV polypeptides and production of virus-like
             particles
JOURNAL      Patent: US 6602705-A 82 05-AUG-2003;
FEATURES     Location/Qualifiers
  source     1..2306
             /organism="unknown"
             /mol_type="genomic DNA"
ORIGIN

Query Match        82.2%;   Score 2019.2;   DB 6;   Length 2306;
Best Local Similarity   93.2%;   Pred. No. 1.5e-202;
Matches 2150;  Conservative   0; Mismatches 138;  Indels  18;  Gaps   3;

Qy    170 GCGCAAGGAGGGCACCAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCCGGC 229
          ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db      1 GCGGGCCGCGAAGGACACCAAATGAAAGATTGCACTGAGAGACCAGCGCTAATTTCTTCCGGG 60

Qy    230 AGGACCTGGCCTTCCCCAGGGCAAGGGCGAGCTGCCAGGTGCGGCAGCAGACCGCGCCA 289
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGGACCTGGCCTTCTCCTGCAGGGCAAGGGCCCGCGAGCTGCAGCAGCAGACCGCGCCA 120

Qy    290 ACAGCCCACCAGCCGCGAGCCGCCAGGGCACCCTG------ACAACCCCGCCAGCGCGAGG 343
          ||||||||||||||||||||||||||||||||||       ||||||||||||||||||||
Db    121 ACAGCCCACCCGCCGCGAGCCGCGCCCGGCGCGGCGGAGAACAACAGCCTGAGCGCGAGG 180

Qy    344 CCGGGCCGAGCGCCAGGGCACCCTG------AACTTCCCCGACGCAAGGGCCCGTGTGGCAGC 397
          |||||||||||||||||||||||||       ||||||||||||||||||||||||||||||
Db    181 CCGGGCCGACCGCCAGGGCACCGTGTGCAAGCTTCCCCGATCAAGGAGGCCCAGCTGGGCAGC 240

Qy    398 GCCCCCTGGTGAGCTGAGAGATGAGCCTGCCCCGCCAGTCAAGGATGGCCCAAGCGCCG 457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCCCCTGGTGACCATCAGGATCGGCGCGGCCAGCTCAAGGAGGCCGCTGCTCGACACCGGCG 300

Qy    458 CCGAGACACACCTGTCTGGAGGAGATGAGCCTTCGCCCGCCAGTGAAGCCCAAGATGATCG 517
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    301 CCGGACGACACCCTGCTGGAGGAGATGAACCTGCCCCCGGCCAGTACGACCCAAGATGATCG 360

Qy    518 GCGGGATCGGCGGCGCTTCATCAAGGTGCGCCAGTCAAGGTGCGGCCAGATCTGCG 577
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCGGGATCGGCGGCGCTTCATCAAGGTGCGCCAGTCAAGGTGCGGCCAGATCTGCG 420

Qy    578 GCAAGAAGGCCATCGCACCGTGCTGATCGGCGACCCACCCCCGTGAACATCATCGGCGCCA 637
          |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCACAAGGCCATCGGCATCGGCGTGCTGTGGGGCCCCACCCCCGTGAACATCATCGGCGCCA 480

Qy    638 ACATGCTGACCCAGCTGGGCATGGACGGCCCGCGGCATGGACTTCCCCATCAGCGCCATCGCCTGC 697
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ACCTGCTGACCCAGATCGGCTGCTGTGAAGCCGAGGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCCGCC 540

Qy    698 CCGTGAAGCTGAAGCCCGGCATGGAAGCCAGTGCCCCTGACCGAGG 757
          |||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCGTGCAGAAGGTGAGCACCGAGGAGGCCTGGGATGGAGTACTGGCCACCTGA 600

Qy    758 AGAAGATCAAGGCCTGACCGCCATCTGCGAGGAGATGGAGGAGGGGCAAGATCACCA 817
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGAAGATCAAGGCTGACCGTGTGGAGGATGGAGGAGGAGGAGGGGCAAGATCAGCA 660

Qy    818 AGATCGGCCCCCAGAACCCTACAACACCCCCTACAACACCCCCGAGAACCGGGCGCCA 877
          ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db    661 AGATCGGCCCCCAGAACCCTCACAACACCCCCGAGAAGAAGAAGAAGGAAGAGGGACGGCAGCA 720

Qy    878 CCAAGTGCGCCAAGCTGGTGAAGCTTCCGCGACTTCGGCGAGCTGGCCCTGCTGAGACTTCTGGG 937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCAAGTGCGCGAAGCTGTGGACTTCGGCGAGCTGGCCCTGGCCGACAACGGCCACCCCAGGACTTCTGGG 780

Qy    938 AGGTGCAGCTGGGCATCCCCGAGAACCTTCGGCTGCTGGACGTGGGACGCGCCTACTTCAGCGACTACACCG 997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGGTGCAGCTGGGCATCCCCGAGAACCTTCGGCTGCTGGACGTGGGGGCGACGCGCCTACTTCAGCGACTACACCG 840

Qy    998 TGGACGTGGGGCGACGCCTACTTCAGGTGCCCCTGGACGACTTCCGCACAGGTACACCG 1057
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGGACGTGGGGCGACGCCTACTTCAGCGACTTCCGCGACAAGGAGTTCCGCCAGTGTGGCCTCCAGCCGCGCCACCG 900
```

```
Qy   1058 CCTTCACCATCCCAGCATCAACAAGCGCATCCGCCAGTACAAGTGC 1117
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CCTTCACCATCCCAGCATCAACAAGCGCATCCGCTACCAGTACAAGTGC 960

Qy   1118 TGCCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGCCCCAGGGCTGGAAGGGCAGCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020

Qy   1178 AGGCCCTTCGCGCCGCAACCCCGAGATCGTGATCATCACCAGGCCCCCCTGTACGTGGGCA 1237
          ||||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGGCCCTTCGCGAAGCAGCAGAACCCCGAGATCGTGATCATCACCAGGCCCCCCTGTACGTGGGCA 1080

Qy   1238 GCGACCTGGAGATCGGCCAGCACCGGCCAAGATCGAGGAGGAGCTGCCGCAAGCAACCTGCTGC 1297
          ||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||
Db   1081 GCGACCTGGAGATCGGCCAGCACCGGCCGCAAGATCGAGGAGGAGCTGCCGCAAGCAACCTGC-- 1140

Qy   1298 GCTGGGGCTTCACACCCCGACAAGAGCACCAGGAGAGCACCGAGAGGAGAGCCCCCTTCCTGCCCAT-- 1355
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1141 GCTGGGGCTTCACACCCCGACAAGCACCAGGAGCACCAGGAGAGCACCCCCCTTCCTGTGATGG 1200

Qy   1356 ---CGAGCTGCACCCGACACAGTGGACCGTGCGCCATCGAGCTGCCCGAGAAGGAGA 1411
             |||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1201 GCTACGAGCTGCACCCGACAGATCCAGAAGTTGGTGGCGCAACGTCGCGCCAGCCAGATCT 1260

Qy   1412 GCTGGACCGTGAACGACATCCAGAGACTTCAGAAGTGTGGGCAAGCTGAACTGGCCAGCCAGATCT 1471
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCTGGACCGTGAACGACATCCAGAAGTTGGTGGCGCAAGCTGAACTGGCCCAGCCAGATCT 1320

Qy   1472 ACCCCGGCATCAAGGTGCGCCAGTCTGTGCAAGGTGCTGCGCGGCGCCAAGGGCCCTGACCG 1531
          |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ACGGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGTGCTGCGCGGCCCACCAAGGCCCTGACCG 1380

Qy   1532 ACATCGTGCCCCTGACCGAGGAGCAAGGTGGCCGAGCTGGCCGAGATCTGC 1591
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 AGGTGATCCCCCTGACCGAGGAGCAAGGTGGCCGAGCTGGCCGAGATCTGA 1440

Qy   1592 GCGAGCCCGTGCACCGGCGTGTACTACGACCCCAGCAACCAAGAAGGAGGACCTGGTGGCCCTTCAAGAACCTGAAGA 1651
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AGGAGCCCGTGCACGAGGTGTACTACGACCCCAGCAACCAAGAAGGAGGACCTGGTGGCCCTTCAAGAACCTGAAGA 1500

Qy   1652 AGCAGGGCCACGACCAGGTGGACTGGACCTACCAGGAGGCCCTTCAAGAACCTGAAGA 1711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AGCAGGGCCAGGGCCAGTGACTGGACCTACCAGGAGGCCCTTCAAGAACCTGAAGA 1560

Qy   1712 CCGGCAAGTACGGCCAAGATGCGACGTGACCGAGGCACACGAGCAGCGTGACCGAGG 1771
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CCGGCAAGTACGGCCAAGATGCCCGCATGGCCCATGCCGAGGCACGAGTGAAGCAGTGACCGAGG 1620

Qy   1772 CCGTGCAGAAGATCGCCATGGACAGCAGTGCATCGTGAATCGTGATCTGGGGGCAAGACCCCCAAGTTCCGCC 1831
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCGTGCAGAAGGTGAGCACCGAGAGCCAGAGCATCGTGAATCGTGATCTGGGGGCAAGCAGAAGTTCAAGC 1680

Qy   1832 TGCCCATCCAGAAGGACCTGCCGAGGAGATGGAGGAGGGGCAAGATCACCA 1891
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGCCCATCCAGAAGGAGGAGAGGGGCTGGGATGGAGTGACTACTGGAGGCCACCTGA 1740

Qy   1892 TCCCCGAGTTGGGAGTTCGTCAGACACCCCCCCCCGCCGCCAAGATCGTGACCTGGGAGA 1951
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TCCCCGAGTTGGGAGTTCGTGAACACCCCCCCCCCGTGGTGAAGCGTGTGACCAGCTGAGA 1800

Qy   1952 AGGAGCCCATCATCGGCGCCGAGACTTCTGGGCCGCACCCCCAGAAGATCGTGAGCCGTGACCTGCGG 2011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGGAGCCCATCATCGGCGGCGCCGGGCCGGGCCATCCAGCTTCTGGGCCACCCCAGGACCA 1860

Qy   2012 AGATCGGCCAGGGCCGCTACGTGACCGCCGGCCGCAGAAGATCGTGAGCCGTGACCTGCCGG 2071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 AGTCGGGCAGGGCCGCTACGTGACCGACCGACAACCGGACCGCGGCCAGAAGGTGTGACATCGCGG 1920

Qy   2072 AGACCACCACCAGAAGACCGAGCTGCGACGAGACTTCCGCAGCGGCA 2131
          ||||  ||||||||||||||||||||||||||||||||||||||||
Db   1921 ACACCACCAACCAGAAGAACCAGCTGCAGGCCATCCACCTGGCCCTGCAGGACAGCGGGGCC 1980

Qy   2132 GCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCG 2191
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT 8
AX427930
LOCUS       AX427930                9166 bp    DNA         linear   PAT 20-JUN-2002
DEFINITION  Sequence 168 from Patent WO0232943.
ACCESSION   AX427930
VERSION     AX427930.1  GI:21538017
KEYWORDS
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 168 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES    Location/Qualifiers
  source    1..9166
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="plasmid pVRI012x/s containing HIV genes"
ORIGIN

Query Match    81.4%;  Score 2000.8;  DB 6;  Length 9166;
Best Local Similarity  91.3%;  Pred. No. 8.8e-201;
Matches 2185;  Conservative  0;  Mismatches  182;  Indels  25;  Gaps  5;
```

```
Qy     14 TGGCCGAGGCCATGAGCCAGGCCACCAGGCCACCAGGCCAGGCCCAACTCCTGATGCAGCGCAGCAACTTCA    73
          |||||||||||||||||||||||||  || |||||||| ||||||||||||||||||||||||||||||||
Db   2961 TGGCCGAGGCCATGAGCCAGGCCACCAGGTGAACACACACCAAACACACACACAGCGCGCAGCAACTGCA   3020

Qy     74 AGGGCCCCAAGGCCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA    133
          ||||||||||||||||||||||||||| |||||||||||| |||||||| || ||||
Db   3021 AGGGCCCCAAGGCCACCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGGCCCGCA   3080

Qy    134 ACTGCCGCGCCCCCGCCCCCGCCCCCGCCCAAGAGGGCTGCTGAAGTGCGGCAAGAAGGGCCACCCAGATGA    193
          |||||||||||||||||||||||||||||  |||||||| |||||||||||||||||| ||||||||||
Db   3081 ACTGCCGCGCCCCCGCCCCCGCCCCCGCCCCCCAAGAGGGCTGCTGAAGTCGCGGCAAGAAGGGCCACCCAGGTGA   3140

Qy    194 AGGACTGCACCGAGCGCCAGGCGACCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA    253
          |||||||||||||| |||||||||||||||||||||||||||||||||| |||||||||||
Db   3141 AGGACTGCACCGAGCGCCAGGCGACCAGGCTAA-TTTTTTAGGGAGATCTGGCCTTCCCCACAAGGGA   3199

Qy    254 AGGGCCCGCGCCCCGAGTTCCCCGGCCAGCAGAACAGCCGCGAGAACCCGGCAAGCAGCGGCCTG    313
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   3200 AGGCCAGGGAATTTTCTTCACCCGAGCTGACCAGGCTAAAGAAGAGAGAGTTC   3259

Qy    314 AGTTGCGCGG------CGACAACCCCGCCGCCAGCGGCCGGCCCGCCCAGGGCA---    364
          ||||||||||||    |||||||||||||
Db   3260 AGGTTTGGGGAAGAGACACAACTCCCTCTCAGAGCAGGCCCCCTTGGTGAGTGAGCTGGATGAGCAGAGGACTG   3319

Qy    365 ---CCCTGAACTTCCCCGAGATCACCGTGTGGCAGATCACCGTCTTTGGCAGCGAACCCCTCGTCACAAGATAG    421
             |||||||||||||||||| ||||||||||||||||||||||||| |||| |||||||| ||||| |||
Db   3320 TATCCTTTAGCTTCCCCAGATCACTCTTTGGCAGCGAACCCTGTTGGTGAGTGAGCTCAAGATAAAGATAG   3379

Qy    422 GCGGCCGAGATCAAGGAGCCGTGGACGCGGCCGCCCACAAGGGCAGATCGCAACAGAGATCGCGACAAGC    481
          ||||||||||||||||||| |||||| ||||| ||||||||||  ||||||||||||||| ||||||
Db   3380 GGGGCCGAGATCAAGGAGCCCTTTCGAGACATCGGCCGCAACAAGGGCACAGTGACTGGACCAAGC   3439

Qy    482 TGGTGGCAAGGCGCATCCGCCAAGGTGCTGTTCCTGACGAGCCATGCGCGGCCATCGTGA    541
          ||||||||||||||||| ||||||||||||||||| ||| |||| |||||||||||||||
Db   3440 TGAACTGCCCGGCCGCGTCAGCCAGCTGGAAGCCCTGGAAGCCCTCATCAAGG   3499

Qy    542 TGCCAGTACGACCAGATTCGGCCCGCGCCCCTGGGCGGCGGCGCGGCCATCGTGATTAAAGC    601
          ||||||||||||| |||||||||| |||| ||||||||||||||||||||||||||||||||
Db   3500 TGGCCAGTACGACCAGATTCGGCCGCGCCCCTGTGGCGGCCACCGGTGC   3559

Qy    602 TGATCGGGCCCCACCCCGTGAACATCATCGGCCCACAATGCTGACCCAGCTGGGCTGCA    661
          |||||||||||||||||||| |||||||||||| ||| |||| ||||||||||||||||||
Db   3560 TGGTTGGGGCCCCACCCCGTGAACATCATCGGCCCGCAACCTGCTGACCCAGCTGGGCTGCA   3619

Qy    662 CCCTGAACTTCCCCATCGGCCCATCGGCGACCGTGCCCGTGGACGTGCCCGTGAAGCCCGCCATCG    721
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   3620 CCCTGAACTTCCCCATCAGCCCCATCAGCGACCGTGCCCGTGACCGTGCCCGTGAAGCCCGCCA   3679

Qy    722 ACGGGCCCCAAGGTGAAGCAGTGCCCCTGACCAGGAGAAGATCAAGGCCCCTGACCGCCA    781
          |||||||||||||||||| |||||||||||||||| |||||||||||||||||||||||
Db   3680 ACGGGCCCCAAGGTGAAGCAGTGCCCCTGACCAGGAGAAGATCAAGGCCCTTGTGTGAGA   3739

Qy    782 TCTCGGAGGATGGAGGAGGGCCAAGTACCACCAAGATCGGCCCCAGAACCCCTACA    841
          ||| |||||||||||||||| |||| | ||||||||| ||||| |||||||||||||
Db   3740 TCTGACCGGATGGAGGAGGGCAAGATCACGCAAGATCGGCCCCGAGAACCCCTACA   3799

Qy    842 ACACCCCCGTGTTGCCATCAAGAGAAGGAGCAGCACCAAGTGCGCAAGTGTGGGACT    901
          ||||||||||||| || ||||||||||||||||||||||||||| ||||| |||||||
Db   3800 ACACCCCCGTGTTCGCACCATCAAGAAGAAGGAGCAGCACCAAGTGCGCAAGTGCTGGGACT   3859

Qy    902 TCCGCCGAGCTGAACAAGCCACCCAGACTTCTGGAGGTGCGAGCTGGCATCCCCACC    961
          |||| |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   3860 TCCGCCGAGCTGAACAAGCCACCCAGACTTCTGGAGGTGCTGACGTGCATCCCCACC   3919

Qy    962 CCGCGGCCTGAAGAAGAGAGCGTGACCGTGGCGCAAGTGTCGTGACGCGCTCGTGGACT   1021
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   3920 CCGCGGCCTGAAGAAGAGCAGCAGAAGAGCGGTGACGTGCTGGACGTGGGGCGACGCCTACTTCA   3979

Qy   1022 GCGTGCCCCTGGACGAGCAGCAGACTTCCGCAGCAGCCTTCACCATCCCCAGCATCAACA   1081
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   3980 GCGTGCCCCTGGACAAGGACTTCCGCAAGTACAACGGGCTGAAGGGCAGCAGCC   4039

Qy   1082 ACGAGACCCCGGCATCGGCTACCAGTACAACGTGCCCCAGGGTTGAAGGGCAGCC   1141
          |||||||||||||||||||||| ||||||||||||| |||||||| ||||||||||
Db   4040 ACGAGACCCCGGCATCGGCTACCAGTACAACGTGCTGCGCGGGCTGAAGGGCAGCAGCC   4099

Qy   1142 CCAGCATCTTCCAGTCGTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA   1201
          ||| |||||||| | || ||||||||||||||||||||||||| ||||||||
Db   4100 CCGCCATCTTCCAGTCGCAGCATGCAGCATGACAAGGCCAAGCAGAACCAGACCCCCGCCA   4159

Qy   1202 AGATCGTGATCTACCA---------GGCCCCCGTGTACGTGGGCAGAGATCGGCC   1255
          |||||||||||||| ||        || ||| ||||||||||||| || ||||||||
Db   4160 ACATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGGTCGGACCTGGAGATCGGCC   4219

Qy   1256 AGCACCGCAAGATCGAGAGCTGCGGCAAGCTGCTGCGTGGGCTTCACCACCC   1315
          ||||||||| |||||||||||||||||| ||||||||| ||||||||||||||
Db   4220 AGCACCGCACCAAGATCGAGGGGCTGCGCGGGCTGCTGCGGGGCTTCACCACCC   4279

Qy   1316 CCGAAGCAGAGAGCTGCACCCCCTTCCGCCCAT-------CGAGCTGCACCCCG   1369
          ||||||||| ||||||||||||||||  |||||       |||||||||||||||
Db   4280 CCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCG   4339

Qy   1370 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGACA   1429
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
VERSION      BD263706.1  GI:33073474
KEYWORDS     JP 2002533124-A/73.
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1  (bases 1 to 2312)
  AUTHORS    Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
             Greer,C., Selby,M. and Walker,C.
  TITLE      Improved expression of HIV polypeptides and production of
             virus-like particles
  JOURNAL    Patent: JP 2002533124-A 73 08-OCT-2002;
             CHIRON CORP
COMMENT      OS   Artificial Sequence
             PN   JP 2002533124-A/73
             PD   08-OCT-2002
             PF   30-DEC-1999 JP 2000591193
             PR   31-DEC-1998 US   60/114495,01-DEC-1999 US   60/168471 PI
             SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
             KARIN HARTOG,
             PI   HONG LIU, CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
             C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
             C12N5/10,
             PC   C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
             Description of Artificial Sequence: FS(-).promod.RTopt(+) FH
             Location/Qualifiers
                  1..2312
FT   source
FT                /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1..2312
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN
Query Match    81.3%;  Score 1998.4;  DB 6;  Length 2312;
Best Local Similarity  92.9%;  Pred. No. 2.3e-200;
Matches 2147;  Conservative  0;  Mismatches 141;  Indels 24;  Gaps 4;
```

```
Qy      170  GCGGCAAGGAGGAGCCACCAGATGAGGACTGCACCGAGGCCCAGGCCAACTTCTTCCGCG  229
             || |||| ||||||||| ||| |||  |||||||||||||||||||||||||  ||||||||
Db        1  GCGGGCCGCGAAGGACACCAAATGAAGAGTGCACTGAGAGAAGATTGCACTCAATTTCTTCCGCG   60

Qy      230  AGGACCTGGCCTTCCCCGGCCAAGGCCAGGCCCGGGAGGCCCAGGGAGTTCCCGAGTTCCCGGCCA  289
             |||||||||||||||  ||||| || |||||||| || |||| || ||||||||||||| |||||
Db       61  AGGACCTGGCCTTCTCTGCAGGCAAGGCCAGGCCCGGGAGTTCAGCAGGAGCCCCGGGCCA  120

Qy      290  ACAGCCCACCAGCCGCGAGCTGCCAGGTGCAGTGCGCGG------ACAACCCCGGCAGCCGGAGG  343
             | |||||||||||||||||||||||| ||||| | |||||      |||  || ||||| |||||||
Db      121  ACAGCCCACCAGCCGCGAGCTGCCAGGTGCAGTGCGCGG------ACAACAACAGCAGCCGGAGG  180

Qy      344  CCGGGCGCGAGGCCCAGGGCCACCCTG-------AACTTCCCCAGATCACCCTGTGGCAGC  397
             ||||||||||||||||| ||| ||||||       ||||||||||| |||||||||||||||
Db      181  CCGGGCGCGAGGCCCAGGGCCACCGTGAGCTTCAAGTTCCCCAGATCAACTTCAACTTCCCCAGATCACCCTGTGGCAGC  240

Qy      398  GCCCCTGGTGAGCATCAAGGTGGGCGGCATCAAGGTGTACCTGAGCTGGGTGC  457
             ||||||||| ||||||||||||||||||| |||||| ||||||||||| ||||||
Db      241  GCCCCCTGGTGACCATCAGGATCGGCGGGCATCAAGGTGTACCTGGGTGGGGTGC  300

Qy      458  CCGACAGCACCGTGCTGGAGGAGATCAGCCTGCCCGGCGCAGTGGAGCCCAAGATGATCG  517
             ||||| ||||||||| |||||| |||| ||||||| ||||| |||||||||||||||||||
Db      301  CCGACAACACCGTGCTGGAGGAGATGAAGCTGGGCGGACGCGTGGAGCGCGGGATCC  360

Qy      518  GCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTGACCAGATCCTGATCGGAGATCTGCG  577
             ||| |||||| |||||| |||||||||||| ||| |||  || | |||||||||| |||||
Db      361  GCGGGATCGGCGGGCTTCATCAAGGTGCGGCCAGTAGGATCGGCGGCATCGGCGGCATCGGCGGCTTCGCG  420

Qy      578  GCAAGAGGCCATCGGCCACCGTGCTGATCGGCCCGGTGGAGCCCACCCCCGGCCCGCGCCA  637
             ||||||||||||||| ||||||| |||| ||| |||||||||| ||||||| |||| |||||
Db      421  GCCACAAGGCCATCGGCCACCGTGGCTGATCGGCCCGTGGGGCCCCACCCCCGGCCCGCGCCA  480

Qy      638  ACATGCTGACCAGCTGGCTGCACCCTGAACTTCCCCATCAGCCCGTGACCACCGTGC  697
             ||||||||||||| |||||| |||| ||||| |||||||||||||||| || ||||||
Db      481  ACCTGCTGACCAGCTGGCTGCACCCTGAACTTCCCCATCAGCCCGTGACCACCGTGC  540
```

```
Db     4340  ACAAGTGGACCGTGCAGCCCATCGTGTGCCCGGAGAAGGACAGCTGGACCGTGAACGACA  4399
Qy     1430  TCCAGAAGCTGTGGGCAAGCTGAACTGGGCCCGGCCAGCCAGATCTACCCGGCATCAAGGTGC  1489
Db     4400  TCCAGAAGCTGTGGGCAAGCTGAACTGGGCCCGGCCAGCCAGATCTACCCGGCATCAAGGTGC  4459
Qy     1490  GCCAGCTGTGCAAGCTGCTGCGGGGCGCACCAAGGCCCTGTCCCCCTGACCG  1549
Db     4460  GCCAGCTGTGCAAGCTGCTGCGGGGCCACCAAGGTGGTGCCCCTGACCG  4519
Qy     1550  AGGAGGGCGAGCTGGAGCTGGCCGAGAACCGGCAGAGATCCTGCGGAGAAGGACCGGGCGCCGGGCCG  1609
Db     4520  AGGAGGGCGAGCTGGAGCTGGCCGAGAACCGGCAGAGATCCTGCGGAGAAGGACCGGGCGCCGGGCCG  4579
Qy     1610  TGTACTACGACCCCAGCAAGAGACCTGTGGCCCGAGATCCAGAAGCAGGAGGGCCACGACCAGT  1669
Db     4580  TGTACTACGACCCCAGCAAGAGACCTGATCGCGCCCGAGATCCAGAAGCAGGAGGGCCACGACCAGT  4639
Qy     1670  GGACCTACCACCAGATCTACCAGGAGGCCCTTCAAGAACCTGAAGACCGGCCAAGTACGCGCCAAGA  1729
Db     4640  GGACCTACCACCAGATCTACCAGGAGGCCCTTCAAGAACCTGAAGACCGGCCAAGTACGCGCCCGCA  4699
Qy     1730  TGCGCACCGCCCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCA  1789
Db     4700  TGAAGGGCGCCCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCA  4759
Qy     1790  TGGAGACCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGA  1849
Db     4760  CCGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAGGAGA  4819
Qy     1850  CCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACTGGATCCCCGAGTGGGAGTTCG  1909
Db     4820  CCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACTGGATCCCCGAGTGGGAGTTCG  4879
Qy     1910  TGAACACCCCCCCCCTGGTGAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG  1969
Db     4880  TGAACACCCCCCCCCTGGTGAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG  4939
Qy     1970  CCGAGACCTTCTACGTGGACGGCCGCCAACCGTGAGACGACCTTCTACGTGGACGGCCCGCT  2029
Db     4940  CCGAGACCTTCTACGTGGACGGCCCGCCAACCGTGAGACCTGAGAACCGCCAGAGGCCGCT  4999
Qy     2030  ACGTGACCAGCACCGGGGCCGCCGCAGAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGA  2089
Db     5000  ACGTGACCAGCACCGGGGCCGCCGCAGAGAGTGGTGCCCCTGACCGACCAACCAACCAGAGAGA  5059
Qy     2090  CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCAGGTGAACATCGTGA  2149
Db     5060  CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCAGGTGAACATCGTGA  5119
Qy     2150  CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGACGGAGC  2209
Db     5120  CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC  5179
Qy     2210  TGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGC  2269
Db     5180  TGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGGCTGGGTGC  5239
Qy     2270  CCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGTCGGTGAGCAAGGGCCATCC  2329
Db     5240  CCGCCCACAAGGGCATCGGCGGCAACGAGCAGATGGACAAGTCGGTGAGCGCCGGCATCC  5299
Qy     2330  GCAAGGTGCTGTTCCTGGACGGCATCGATGGCCGGCATCGTGATCTACCAGTA  2381
Db     5300  GCAAGGTGCTGCTTCCTGGACGGCATCGATGGCGGCCAAGGCCGACCAGCACCGAGAA  5351
```

```
RESULT 9
BD263706
LOCUS       BD263706                2312 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Improved expression of HIV polypeptides and production of
            virus-like particles.
ACCESSION   BD263706
```

```
Qy   1766 CCGAGGCCGTGCAGAAGAGAGATCGCCATGCGAGAGCATCGTGATCTGGGCGAAGACCCCCAAGT 1825
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCGAGGCCGTGCAGAAGGTCAGCACCGAGAGCATCGTGATCTGGGCGAAGATCCCCAAGT 1680

Qy   1826 TCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCACTGGCCAGGCCA 1885
          |||||||||||||||||||||||||||||||||||||||||||
Db   1681 TCAAGCTGCCCATCCAGAAGGAGGAGACCTGGGAGGACCTGGTGATGAGTACTGGCCAGGCCA 1740

Qy   1886 CCTGGATCCCCGGAGTGGGAGTTCGTGAACACACCCCCCGTGTGAAGCTGTGGTACCAGC 1945
          |||||||||||||||||||||||||||||||||
Db   1741 CCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCTGTGTGAAGCTGTGGTACCAGC 1800

Qy   1946 TGGAGAAGGAGGCCCATCATCGGCCGAGACCTTCTACGTGGACGCGCCGCCACCGCG 2005
          |||||||||||||||||||||||||||||||||||
Db   1801 TGGAGAAGGAGGCCCATCGTGGGCGCGCCGCCGAGACCTTCTACGTGGACGCGCCCCAACGCG 1860

Qy   2006 AGACCAAGATCCGCAAGGCCGGCTAGGTGACCGTGACGCCGGCGGCGCAGAAGATCGTGAGCC 2065
          |||||||||||||||||||||||||||||||||||||||
Db   1861 AGACCAAGCTGGGCAAGGCCGGCTAGGTGACCGTGACGCCGGCGGCGAAGAAGGTGGTGAGCA 1920

Qy   2066 TGACCGAGACCACCAACCAGAGACCCAGCTGCAGGCGCATCCAGCTGCCTGCAGGACA 2125
          |||||||||||||||||||||||||||||||||||
Db   1921 TCGCCGACACCACCACCAGACCGAGCTGCAGGCGCATCCACCTGCCCTGCAGGACA 1980

Qy   2126 GCGGCAGCGAGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGCCC 2185
          |||||||||||||||||||||||||||||||||
Db   1981 GCGGCCTGGAGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGCCC 2040

Qy   2186 AGCCGGCACAAGAGCGAGAGGAGCGAGCTGTGACCACGATCATCGAGCTGATCAAGAAGG 2245
          |||||||||||||||||||||||||||||||||||
Db   2041 AGCCGGACAAGAGCGAGAGCGAGCTGGTGACCGAGATCATCGAGCTGATCAAGAAGG 2100

Qy   2246 AGAAGGTGTACCTGAGCTGGGTGCCGGCCGCCCAGGCCGCAGATCG 2305
          |||||||||||||||||||||||||||||||||||
Db   2101 AGAAGGTGTACCTGGGCCTGGGCGTGCCCGGCCGCCCAGGCCGCAGCAGGTGG 2160

Qy   2306 ACAAGCTGGTGAGCAAGGACGCTGGTCCTGGCGATCCGAAGGTCGTGTTCCTGGACGACCATGGGCGGCA 2365
          ||||||||||||||||||||||||||||||||
Db   2161 ACAAGCTGGTGAGCAAGGACGCTGGTCCTGGCGATCCGAAGGTGCTGTTCCTGGACGACCATGGGCGGCA 2220

Qy   2366 TCGTGATCTACCAGTACATGGACGACAGCCGGGCCGGCGGCGGCCTAGGATCGATT 2425
          ||||||||||||||||||||||||||||||
Db   2221 TCGTGATCTACCAGTACATGGACGACAGCCGGGCCGGCGGCGGCCCCTAGGATCGATT 2280

Qy   2426 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2457
          ||||||||||||||||||||||||||
Db   2281 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2312
```

```
Qy    698 CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAGCAGTGGCCCCTGACCGAGG 757
          |||||||||||||||||||||||||||||||||||
Db    541 CCGTGAAGCTGAAGCCCGGCGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCCCTGACCGAGG 600

Qy    758 AGAAGATCAAGGCCCTGACCGTGGTGAGAGATCGCGACATCTGCGAGATCTGCACCCCGTGTGAGGCAAGATCACCA 817
          |||||||||||||||||||||||||||||||||||
Db    601 AGAAGATCAAGGCCCTGACCGTGGTGAGAGATCGAGATGGAGAAGGAGGGCAAGATCACGCA 660

Qy    818 AGATCGGCCCCGAGAACCCTACAACACCCCGTGTTCGCCATCAAGAAGGACAGCA 877
          |||||||||||||||||||||||||||||||||||
Db    661 AGATCGGCCCCGAGAACCCTACAACACCCCGTGTTCGCCATCAAGAAGGACAGCA 720

Qy    878 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGGCGACCTGGG 937
          |||||||||||||||||||||||||||||||||||
Db    721 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCGACCTTCTGGG 780

Qy    938 AGGTGCAGTGGCGATCCCCACCCCGCGCGGCCGGCGTGACCGTGC 997
          |||||||||||||||||||||||||||||||||||
Db    781 AGGTGCAGTGGCGATCCCCACCCCGCGCCGGCGTGACCGTGC 840

Qy    998 TGGACGTGGGCGACCCTACTTCAGCGTGCCCCTGACCGTGCCCCTGACCTTCCGCCAGTACACACCG 1057
          |||||||||||||||||||||||||||||||||||
Db    841 TGGACGTGGGCGACCCTACTTCAGCGTGCCCCTGACCGTGCCCCTGACCTTCCGCCAGTACACCG 900

Qy   1058 CCTTCACCATCCCAGCATCAACAACGAGACCCCCGCATCCGCTACCAGTACACAACGTGC 1117
          |||||||||||||||||||||||||||||||||||
Db    901 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGCATCCGCTACCAGTACACAACGTGC 960

Qy   1118 TGCCCCAGGCTGGAGGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
          |||||||||||||||||||||||||||||||||||
Db    961 TGCCCCAGGCTGGAGGGGCTGGAGGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020

Qy   1178 AGCCCTTCCGCGCCCCGAGATCGTGATCTACCA------GGCCCCCTGTACG 1231
          |||||||||||||||||||||||||||||||||||
Db   1021 AGCCCTTCCGCGCAAGCAGAGACCCGGACATCGTGATCTGATCTACCAGTACATGGACGACCTGTGTACG 1080

Qy   1232 TGGGGCAGCAGCAGCCAGTGAGATCGAGAGCGTGCGCAAGCACC 1291
          |||||||||||||||||||||||||||||||||||
Db   1081 TGGGGCAGCAGCAGCCAGTGAGATCGAGAGCGTGCGCAAGCACC 1140

Qy   1292 TGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCCAGAGAGGAGCCCCCTTCCTGC 1351
          |||||||||||||||||||||||||||||||||||
Db   1141 TGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCCAGAGAGGAGCCCCCTTCCTGT 1200

Qy   1352 CCAT------CGAGCTGCACCCCGACCAAGTGACCGTGCAGCCCATCGAGCTGCCCGGAGA 1405
          ||||
Db   1201 GGATGGGCCTGCACGAGCTGCACCCCGACCAAGTGGACCGTGCAGCCCATCGCGCCCGGAGA 1260

Qy   1406 AGGAGAGCTGAACGAGCATCCAGAGACTGGTGCAGCTGGTCGGGCAGCC 1465
          |||||||||||||||||||||||||||||||||||
Db   1261 AGGACAGCTGAACGACATCAGCAGCATCCAGAGACTGGTGTCAGCTGAACTGGGCGGGCCAGCC 1320

Qy   1466 AGATCTACCCCGGCCATCAAGGTGCGCCAGCTGCTGCGCGGCCGGCGCGCCAAGGCCC 1525
          |||||||||||||||||||||||||||||||||||
Db   1321 AGATCTACCCCGGCCATCAAGGTGCTGCAAGTGTGCTGCGCGGCGGCGGCCACCAAGGCCC 1380

Qy   1526 TGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGA 1585
          |||||||||||||||||||||||||||||||||||
Db   1381 TGACCGAGGTGATCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGA 1440

Qy   1586 TCCTGCGGAGCCGGCCGTGCACGGCGTGCTGCTACGACCCGTGGCCGCGGACCTGGTGCCGGGA 1645
          |||||||||||||||||||||||||||||||||||
Db   1441 TCCTGAAGGAGCCGGCCGTGCACGGCGTGCTGCTACGACCCGTGGCCGCGGACCTGGTGCCGGGA 1500

Qy   1646 TCCAGAAGCAGGGCCACGGCCAGTGACCTACCACCGATCTACCAGCCCTTCAAGAACC 1705
          |||||||||||||||||||||||||||||||||||
Db   1501 TCCAGAAGCAGGGCCAGGGCCAGTGACCTACCACCGATCTACCAGCCCTTCAAGAACC 1560

Qy   1706 TGAAGACCGGCAAGTACGCCAAGATGCGCAACGCCCACACCAACCGCGGCGCTGA 1765
          |||||||||||||||||||||||||||||||||||
Db   1561 TGAAGACCGGCAAGTACGCCAAGATGCGCAACGCCCACACCAACCGCGGCGCTGA 1620
```

```
RESULT 10
AR373389
LOCUS       AR373389     2312 bp    DNA     linear    PAT 18-DEC-2003
DEFINITION  Sequence 84 from patent US 6602705.
ACCESSION   AR373389
VERSION     AR373389.1  GI:40075492
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2312)
  AUTHORS   Barnett,S.W., Megede,J., Greer,C. and Selby,M.
  TITLE     Expression of HIV polypeptides and production of virus-like
            particles
  JOURNAL   Patent: US 6602705-A 84 05-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..2312
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match      81.3%;   Score 1998.4;   DB 6;   Length 2312;
Best Local Similarity  92.9%;
Matches 2147;  Conservative 0;  Mismatches 141;  Indels 24;  Gaps 4;
```

Qy 1232 TGGGCAGCGACCTGAGGATCGGCAGCACCGCGCCGCACCGGTGCCAAGCACC 1291
          ||||| |||||||| |||||||||||||||| |||||| |||||| ||||||
Db 1081 TGGGCGACCGACCTGGAGATCGGCAGCACCGCACCGCACCCAAGATCGAGGAGCTGGGCCAGCACC 1140

Qy 1292 TGCTGGCGTGGGGCTTCACCACCGGTGCCAGCGAGCGCAAGGCGCCCCCTTCCTGC 1351
          ||||| ||||||||||||||||||||||||| |||||||||||||||||||||
Db 1141 TGCTGCGCTGGGGCTTCACCACCGGTGCCAGCGACCGCAAGGAGCCCCCTTCCTGT 1200

Qy 1352 CCAT------CGGAGTGCACCGGACCAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGA 1405
                    ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 GGATGGGCTACAGCGGAGTGCACCGGACCAAGTGGACCGTGCAGCCCATCGTGCTGCCCGAGA 1260

Qy 1406 AGGAGAGCTGGACCGTGAACGACACATCCAGAGAGCTCGTGTGGCAAGCTCGGGCCAGCC 1465
          |||||||| |||||||||||| ||||| |||||| |||||||||||||||||||||||||
Db 1261 AGGACAGCTGGACCGTGAACGACATCCAGAGAGCTCGTGGTGGCAAGCTGGGCCAGCC 1320

Qy 1466 AGATTCACCCCGGCATCAAGTGCGCCAGCTGTCGCAAGCTGTCGCGGCGGCCAAGCGCCC 1525
          |||| ||||||||||||||||||| ||||||||||||| |||||||||||||||||||||
Db 1321 AGATCTACGCCGGCATCAAGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCCACCAGGGCCC 1380

Qy 1526 TGACCGACATGTGCCCTGACCGAGGAGGAGCTGAGCTGGCCGAGAACCGCGGAGA 1585
          ||||||||||||||||||| ||||||||||||||||||||||||||||| |||||
Db 1381 TGACCGAGGTGATCCCCGGCTGACCCGAGGAGCTGAGCTGGCCGAGAACCGCGGAGA 1440

Qy 1586 TCCTGCCGCGAGCCGTGCACGGCGTGTACTACGACACCCCAGCAAGGACCTTGGTGCCGAGA 1645
          |||||||||||||||||||||||||||| |||||||||||| ||||||||||||||||||
Db 1441 TCCTGAAGGAGCCCGTGCACGAGGTGTACTACGACACCCCAGCAAGGACCTGGTGCCGAGA 1500

Qy 1646 TCCAGAAGCAGGGCCACGACCAGTGACCTGGACCTACCACGGAGCCTTGAAGCAGCTGA 1705
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db 1501 TCCAGAAGCAGGGCCAGTGACCTGGACCTACCAGATCTACCAGGAGCCTTCAAGACC 1560

Qy 1706 TGAAGACCGGCAAGTACGCCAAGATGCCCAAGACGTCGTGATCTCGGGGCAAGACCCCCAAGT 1765
          ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db 1561 TGAAGACCGGCAAGTACGCCCATGCCCATGCAGCAGCCGGATCGTGATCTCGGGGCAAGATCCCCAAGT 1620

Qy 1766 CCGAGGCCGTGCAGAAGATCGCCATGGACACCGGAGAGCATCGTGATCGGGACCGAGG 1825
          |||||||||||||||||||| ||||||||||||| |||||||||||||||||||||||
Db 1621 CCGAGGCCGTGCAGAAGGTGAGCACCAGCGAGACCTGGTGGAACGTGATCTCGGGCAAGATCCCCAAGT 1680

Qy 1826 TCGGCTCCCATCCAGAAGGAGAGACCTGGTGGACCGACTACTGGCAGGCCA 1885
          ||| ||||||||||||||||||||||||||| || ||||| ||||||||
Db 1681 TCAAGTGCCCATCCAGAAGGAGACCTGGGATGGATGGCTACTGGCAGGCCA 1740

Qy 1886 CCTGGATCCCCGAGTGGGGAGTTCGTGAACACCCCGTGAAGGTGTGTGTGTACCAGC 1945
          |||| |||||||||||||||||||||| ||||||||||||||||||||||||||||
Db 1741 CCTGGACACCCCGAGTGGGGAGTTCGTGAACACCCCGTGGTGAAGGTGTGTGGTACCAGC 1800

Qy 1946 TGGAGAAGGAGCCCATCATGGCGGCCGCGACCTTCTACGTGCGACGGCGCGCCAACCCG 2005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 TGGAGAAGGAGCCCATCGTGTGGCGGCCGCGACCTTCTACGTGCGACGGCGCGCCAACCCG 1860

Qy 2006 AGACCAAGATCGGCAAGGCCGGCTTACGTGACCTGGGGCCGCAGAAGATCGTGAGCC 2065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 AGACCAAGATCGGCAAGGCCGGCTTACGTGACCTGGGGCCGCAGAAGGTGTGTGAGCA 1920

Qy 2066 TGACCGAGAACCAACCAGAAGACCGAGCTGCGAGGCCATCCAGCTGCCCCTGCAGCAGACA 2125
          ||||| |||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 1921 TCGCCGACACCAACCAGAAGACCGAGCTGCGAGGCCATCCACCTGCCCCTGCAGCAGACA 1980

Qy 2126 GCGGCAGCGAGGTGAACATCGTGACGCAGCCAGTACGGCCCTTGGGCATCATCCAGGCCCCG 2185
          |||||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db 1981 GCGGCCTGGAGGTGAACATCGTGACGCAGCCAGTACGGCCCTGGGCATCATCCAGGCCCCG 2040

Qy 2186 AGCCCGACAAGAGCCGAGAGCGAGTCGGTGAACCAGCAGCTGAGCAGCTGATCAAGAGAGG 2245
          ||||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db 2041 AGCCCGACAACAAGAGCCGAGAGCGAGTCGGTGAACCAGCAGCTGAGCAGCTGATCAAGAAGG 2100

Qy 2246 AGAAGGGTGTACCTGACGTGGGTGCCCGCCACAAGGGCGCAACGAGCGATCG 2305
          |||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db 2101 AGAAGGGTGTACCTGGCCTGGCCTGGGTGCCCCCAAGGGCATCGGGCGGGCAACGGCAGGTGG 2160

Qy 2306 ACAAGTCTGGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGTGATGGCGGCA 2365

Qy 170   GCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGGCCAACTTCTTCCGCG 229
          |||||||||||| || |||| |||||||||||||| ||||||||||||||||||||
Db 1   GCGGCCGGCGAAGGACACACCACCGGGCCTGAGACTGACTGAGACAGCGGCTAATTTCTTCCGCG 60

Qy 230   AGGACCTGGCCTTCCCGCGAGTTCCCAGCGAGCGCAAGGCCCGCGAACCGCGGCCA 289
          |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||
Db 61   AGGACCTGGCCTTCCTGCGAGGGCAAGGCCCGCGAGTTCAGCAGCGAGCAGACCCGCGCCA 120

Qy 290   ACAGCCCCACCAGCGCGGAGCGTGCAGGTGCGCGGCG-----ACAACCCCGCCAGCGGAGG 343
          |||||||||||||||||| |||||| |||||||| |||
Db 121   ACAGCCCCACCACCCGCGCGAGCTGCAGGTGTGGGCGCGAGAAACAGCCTGAGCGGAGG 180

Qy 344   CCGGCGCCGAGCGGCCAGGCGCACCCTG------AACTTCCCCAGATCACCCTGTCGCAGC 397
          |||||||||||||||| ||||||||||| |||||||||||||||||| |||||||
Db 181   CCGGCGCCGAGCGGCCAGGGCACCCTGGAGCTTCAACTTCCCCAGATCACCCTGTGGCAGC 240

Qy 398   GCCCCTGGTGCCAGTCATCAAGGTGGGGCGCCAGATGAGCCTGCTCGACACACCGGGCG 457
          ||||||||||||||||||||||||||||| |||||||||| |||||| ||| ||||||
Db 241   GCCCCTGGTGACCATCAAGGTCGGCGGGCCCAGCTCAAGGAGGCGCGTGCTGCTCGACACCGGGCCC 300

Qy 458   CCGACGCACCGTGCTGGAGGAGATGAGCCTGTCCCCGGCAAGTGGAAGCCCAAGATGATCG 517
          |||||||||||||||||||||||||| ||||||||||||||||||||| ||||||||||||
Db 301   CCGACGCACCGTGCTGGAGGAGATGAACGTGTGCCCCGGCAAGTGGAAGCCCAAGATGATCGG 360

Qy 518   GCGGCATCGGGCGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCG 577
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db 361   GCGGGATCGGGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCCCGTGGGAGATCTGCGC 420

Qy 578   GCAAGAAGGCCATCGGACCACGGTGCTGATCGGCCCACCACCCCCGTGAACATCATCGGCCGCA 637
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db 421   GCCACAGGCCCATCGGCGGCTGCACCGTGCTGATCGGCCCACCACCCCCGTGAACATCATCGGCCGCA 480

Qy 638   ACATGCTGACCCAGTCTGGGCCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGC 697
          |||| |||||||||| ||||||||||||||||||||||||| |||||||||||||||||||
Db 481   ACCTGCTGACCCAGATCGGCTGCACCGTGCTGGAGACCCTGAACTTCCCCATCAGCCCCATCGAGACCGGTC 540

Qy 698   CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGCCCCTGACCGAGG 757
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db 541   CCGTGAAGCTGAAGCCCGGGATGGACGGCCCGGGTGTGAAGCAGTCAAGGTCCCCCCTGACCGAGG 600

Qy 758   AGAAGATCAAGGCCCTGACCGACCCATTCTGCGAGGATGGAGAAGGGGCAAGATCACCA 817
          |||||||||| ||||||||||||||||||||||||||||||| ||||||||||||||||
Db 601   AGAAGATCAAGGCCCTGACCGACCCCTGTGGGAGATGGAGGAGAAGGGGACAAGATCAGCA 660

Qy 818   AGATCGGCCCCGAGAACCCTACAACACCCCGTGTTCGGCCATCCAAGAAGAAGACACGCA 877
          |||||||||| ||||||||||||||||| |||||||||||||||||||| |||||||||
Db 661   AGATGCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACACAGCA 720

Qy 878   CCAAGTGGCGCAAGCTGGTGACTTCCGCAGACTTGAACAAGCGCACCCAGGACTTCTGGG 937
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 721   CCAAGTGCGCAAGCTGGTGGACTTCCGCAGACTGAACAAGCGCACCCAGGACTTCTGGG 780

Qy 938   AGGTGCAGCTGGGCATCCCCACCCCCGCCGGCTTGAAGAAGAGGAAGCGGTGACCGTGC 997
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db 781   AGGTGCAGCTGGGCATCCCCCACCCCCGCGGCTTGAAGAAGAGGAAGCGGTGACCGTGC 840

Qy 998   TGGACCTGTGGGGACGACTTCAGCCTGCCCCTGACGTGCCCATCAGCTTCGGCAAGTACACCG 1057
          |||||||| |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 841   TGGACGTGTGGGGCACGGCCTACTTCAGCCTGCCCCTGACGTGCCCATCAAGAAGAAGGACACCG 900

Qy 1058 CCTTCACCATCCCCAGCATCAACACGAGACCCGGCATCTTCCAGAGCACAGATCGTGC 1117
          |||||||||||||||||||||||||||| || ||||||| |||||||||||||||||
Db 901   CCTTCACCATCCCCAGCATCAACAACCAGAGACCCCGGCCATCTTCCAGAGCACAGATCGTGC 960

Qy 1118 TGCCCCAGGGGCTGGGAAGGGCAGCACCCCCCGAGATCTTCCAGAGCATGACCAAGATCCTG 1177
          |||||| |||||||||| ||||||||||||||||||||||||||||||||| ||||||||||
Db 961   TGCCCCAGGGCTGGGAAGGGCAGCACCCCCCGCCATCTTCCAGAGCAGCAGATCCCTGG 1020

Qy 1178 AGCCCTTCCGCGCCCGAGATCGTGATCTTACCA------GGCCCCCTGTACG 1231
          |||||||||||||||||||||||||||||||||| |||||||||||||
Db 1021 AGCCCTTCCGCAAGCAGCAGACAGACCCCGACATCGTGATCTTACCAGTGATCTACCAGTACATGGACGACCTGTACG 1080

```
Db   2161 ACAAGCTGGTGAGCGGCCGGCATCCGCAAGGTGCTCGTTCCTGAACCGGCATCGATGGCGGCA 2220
          |||||||||||||||||    |||||||             ||||||||||||||||||||
Db   3435 TGAACCTGCCCGGCCGCTGGAGCCCGGCATCCGGCGCTTCATCAAGG 3494
          ||||||||||||||

Qy   2366 TCGTGATCTACCAGTACATGACGACCTGTACGTGGCAGCGCCCTAGGATCGATT 2425
          |||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db   2221 TCGTGATCTACCAGTACATGACGACCTGTACGTGGCAGCGCCGCGGCGGCGCTAGGATCGATT 2280
          TGCGCCAGTACGACCAGATCCGATCCGATGAGATCTCCGGCAGGAGGCATCGGCACCGTGC 601
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3495 TGGGCCAGTACGACCAGATCCGATCCGATGAGATCTGGGCCCACAAGGCCATCGGCACCGTGC 3554

Qy   2426 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2457
          |||||||||||||||||||||||||||||||
Qy    602 TGATCGGCCCCGCCCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGCTGGGCTGCA 661
          |||||||||||          |||||||||||||||||||||||||||||||||||||||
Db   2281 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2312
          TGGTGGGCCCACCCCCGTGAACATCATCGGCCGCCAACCTGCTGACCCAGATCGGCTGCA 3614
```

RESULT 11
AX427936

```
Query Match            81.1%;   Score 1992.4;   DB 6;   Length 9788;
Best Local Similarity  91.2%;   Pred. No. 6.5e-200;
Matches  2181;   Conservative   0;   Mismatches  181;   Indels   30;   Gaps   5;

Qy     14 TGGCCGAGGCCATGAGCGAGGCCAGCGCCAGCCCCAGCGCCAGCATCCTGATGCAGCCAGCAACTTCA 73
          |||||||||||||||||||||||||||||||||||||   ||||||| ||||||||||||||||||
Db   2961 TGGCCCAGGCCATGAGCGAGGCCAGCGCCAGCCCCAGCGCCAGCATCATGATGCAGCCAGCAACTGCA 3020

Qy     74 AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 133
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   3021 AGGGCCCCAAGGCCACCACCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGCCCCGCA 3080

Qy    134 ACTGCGGCGCCCCCCGCAAGAAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
          ||||||||||||||||||||||||||||||||||||||| |||||||
Db   3081 ACTGCCGCGCCCCCCGCAAGAAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGTGA 3140

Qy    194 AGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGCGAGAACTTC------TAGGGAAGATCTGGCCCTTCCCCAGGGA 253
          |||||||||||||||||||||||||||||||||   ||    |||||||||||||||||||||
Db   3141 AGGACTGCACCGAGCGGCAGGCTAA------TAGGGAAGATCTGGCCCTTCCCCACCAAGGGA 3194

Qy    254 AGGCCCGCGAGTTCCCCAGGAGCAGAACCGCGCCAACAGCCGCCCCACCGCCGGCGAGCTGC 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3195 AGGCCCAGGAATTTTCTTCAGAGCAGACCAACAACTCCCTTCAGAGCAGACCAGCCAGAGAGAGCTTC 3254

Qy    314 --------CGACACACCCCGAGCAGCGGCCCCGCGAGCGGCCCGGCGCCGCCAGGGCA--- 364
          |||   |||||||||||||||||||||||||||||||||||||||||||||
Db   3255 AGGGCCAGGAATTTTCTTCAGAGCAGACCAACAACTCCCTTCAGAGCAGACCAGCCAGAGAGAGCTTC 3314

Qy    365 ---CCCTGAACTTCCCTCAGATCACCCTCGTGGCAGCGCACCGGCCACCGTGCTCAAGGTGG 421
          |||||||||||||||||||||||||   |||||
Db   3315 TATCCTTCCTTTAGCTTCCCTCAGATCACTTCTTTGGCAGCGACACCGTGCTCAATAAAGATAG 3374

Qy    422 GCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGCCGACGACACCGTGCTCGAGGAGA 481
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3375 GGGGCCAGTCGAAGGAGGCCCTTCAAGGCCCCGGCGCCGCCGACGACACCGTGCTCGAGGAGA 3434

Qy    482 TGAGCCCGTCGCCCGGCAAGTGAGCGGCCGGCATCCGGCAAGATGATCGGCGGCCGCTTCATCAAGG 541
          |||||||||||||||||
Db   1609 AGGAGGCCGAGCTGGAGCTGGCCGAGAATCCGGCCGAGAAGATCGGCGGCATCGGCGCCGCCGTGCACGGCG 1609
```

```
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
                    /note="plasmid pVR1012x/s containing HIV genes"

ORIGIN

          Query Match    80.5%;    Score 1978.4; DB 6; Length 9169;
          Best Local Similarity 91.1%; Pred. No. 1.9e-198;
          Matches 2178; Conservative 0; Mismatches 186; Indels 28; Gaps 6;

Qy     14 TGGCGGAGCCATGAGCCAGGCCAACATCTGATGCAGCGGCAGCAACTTCA 73
          |||| ||||| ||||||||||||||||            |||||||| |||||
Db   2967 TGGGCGGAGGCATGAGCCAGGCCAACCACCACCATGATGCAGCGGCAACTTCA 3026

Qy     74 AGGGCCCCAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCAGGAGA 133
          |||||||||||          ||||||||||||||||||||||||||||||||||
Db   3027 AGGGCCAGAAGCGCAT---CAAGTGCTTCAACTGCGGCAAGGAGGGCCACCAGGAGA 3083

Qy    134 ACTGCGGCGCCCCCCGCCAAGAGAGGCTGTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
          |||||||||||              |||||||||||||||||||||||||||||||||||
Db   3084 ACTGCGGCGCCCCTGCGCAAGAGAGGCTGTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 3143

Qy    194 AGGACTGCACGGAGCCCCAGGCCAACTTCTTCCCGGCCCAACCGCGCGGAGCCGCGAGCTGC 253
          |||||||||||||||||||||||||||        |||||||||||||||||||||||||
Db   3144 AGGACTGCACGGAGCCGACCAGGCTRA-TTTTTAGGGAAGATCTGGCCTTCCCACCAAGGGA 3202

Qy    254 AGGGCCCGGCGAGTTCCCCAGGCCAGCAGAAACCGGCCAACAGCCCACGCGCGAGCTGC 313
          |||||||||              ||||||||||||||||||||||||||||||||||||
Db   3203 AGGGCCAGGGAATTTCTTCAGCAGCCAACCAACAGCCCACCAGAAGAGAGCTTC 3262

Qy    314 AGGTGCGCGG------CGACAACCCCGAGCGGGCCCAGGCAGCAGGCCAGGGCA---- 364
          |||||||||        ||||||||||||||||||||||||||||||||||||
Db   3263 AGGTTTGGGAAGAACAACACTCCCTCAGAAGCAGGAGCCGATAGACAAGGAACTG 3322

Qy    365 ---CCCTGAACTTCCCCAGATCACCGTGGCCAGGGCATCGAGTCAAGGTGG 421
          |||              |||||||||||||||||||||||||||||||||||
Db   3323 TATCCTTTAGCTTCCCTCAGATCACTCTTTGCAGCGACCCTCGTCAGCAATAAAGATAG 3382

Qy    422 GCGGCCAGATCAAGGAGCCCGCCCAGCCGACAAGAGCGAGAGCGAGC 481
          |||||||||||              ||||||||||||||||||||||
Db   3383 GGGGCCAGTGGCCCTTCAGATCATCCGAGGCCGACAAGAGCGAGAGCGAGC 3442

Qy    482 TGAGCCTGCCCGGCAAGTGGAAGCCAAGATGATGCGGCGGCCATCGGCGGCGGCTTCATCAAGG 541
          |||||||              ||||||||||||||||||||||||||||||||||||||||||
Db   3443 TGAACCTGCCCGGCAAGTGGAAGCCCAAGATGATGCGGCGGCCATCGGCGGCGGCTTCATCAAGG 3502

Qy    542 TGCGCCAGTACGACCAGATCCTGATGGAGTCGGACGGGCATCC 601
          ||||||||||||||||||||||||||||||||||||||||||||
Db   3503 TGGGCCAGTACGACCAGATCCTGATGGAGTCGGACGGGCATCC 3562

Qy    602 TGATCGGCGGCCCACCCCATCGAGCACATCGGCCGCCCTGAAGCTGAAGCTGAAGGCCGGCATGG 661
          |||              ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3563 TGGTGGGCCCACCCCATCGAGCACATCGGCCGCCCTGAAGCTGAAGCTGAAGGCCGGCATGG 3622

Qy    662 CCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCCCCTGACCGGCC 721
          ||||              ||||||||||||||||||||||||||||||||||||||||
Db   3623 CCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCCCCTGACCGGCCA 3682

Qy    722 ACGGCCCCAAGGTGAAGCAGTGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCGCCA 781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3683 ACGGCCCCAAGGTGAAGCAGTGCCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGTGAGA 3742

Qy    782 TCTTGCGAGGATGGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACA 841
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3743 TCTTCACCCAGATGGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACA 3802

Qy    842 ACACCCCGTCTGCCATCAAGAAGAAGGACAGCACCAAGTGGCGAAGCTGGTGGACT 901
          |||||||              ||||||||||||||||||||||||||||||||||||
Db   3803 ACACCCCGTTGTCGCCATCAAGAAGAAGGAAGGAGCAGCACCAAGTGGCGCAAGCTGGTGGACT 3862

Qy    902 TCCGCGAGCTGAACAAGGCGCACACCCCAGACTTCTGGGAGGTGCAGCTTGGGCATCCCCACC 961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3863 TCCGCGAGCTGAACAAGGCGCACACCCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC 3922
```

```
Db   4515 AGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGAGCCGTGCACGGCG 4574

Qy   1610 TGTACTACGACCCCAGCAAGAACCTGGTGCCGAGATCCAGAGATCCAGAGGGCCACGACCAGT 1669
          ||||||||||||||||||||              ||||||||||||||||||||||||||||
Db   4575 TGTACTACGACCCCAGCAAGGACCTGATCGGCCAGATCCAGAGATCCAGAGAGCCAGGGCCAGT 4634

Qy   1670 GGACCTACCAGATCTACCAGAGCCCTTCAGAACCTGAAGAACCTGAAGA 1729
          |||||||||||||||||||||||||||              |||||||||
Db   4635 GGACCTACCAGATCTACCAGAGCCCTTCAGAACCTGAAGACCTGAAGACCTACGCCCGCA 4694

Qy   1730 TGCGCACCGCCACCACCAAGCAGTGAAGCAGCAGCTGAAGCCGGCCGTGCAGAGAGATCGCCA 1789
          ||||||||||              |||||||||||||||||||||||||||||||||||||||
Db   4695 TGAAGGGCGCCACCACCAAGCAGTGAAGCAGCAGCTGAAGCCGGCCGTGCAGAGAGATCGCCA 4754

Qy   1790 TGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGA 1849
          |||||||||||||||||||||              ||||||||||||||||||||||||||
Db   4755 CCCAGAGCAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAGGAGA 4814

Qy   1850 CCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCG 1909
          |||||||||||||||||||||||||||              ||||||||||||||||||||
Db   4815 CCTGGGAGGCCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCG 4874

Qy   1910 TGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG 1969
          ||||||||||||||||              |||||||||||||||||||||||||||||||
Db   4875 TGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG 4934

Qy   1970 CCGAGACTTCTACGTGACGGGCGCCGCCAACCGAGATCGTGACCGACCACCAACCAGAAGA 2029
          ||||||||||||||||||||              |||||||||||||||||||||||||||
Db   4935 CCGAGACTTCTACGTGACGGGCGCCGCCAACCGCGAGATCGTGACCGACCACCAACCAGAAGA 4994

Qy   2030 ACGTGACCGACCGGGGCCCGCGAGAGATCGTGACCTGACCGAGACCACCAGAAGA 2089
          ||||||||||||||||||              |||||||||||||||||||||||
Db   4995 ACGTGACCGACCGGGGCCCGCGAGAGGTGGTGCCCGATAGCCTGACCGACCACCAGAAGA 5054

Qy   2090 CCGAGCTGCAGGCCATCCAGCTGCCCTGCAGGACAGCGGCAGCGGGTGAACATCGTGA 2149
          ||||||||||||||              |||||||||||||||||||||||||||||||
Db   5055 CCGAGCTGCAGGCCATCCAGCTGCCCTGCAGGACAGCGGCAGCGGGTGAACATCGTGA 5114

Qy   2150 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC 2209
          |||||||||||||||||||              ||||||||||||||||||||||||||||
Db   5115 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC 5174

Qy   2210 TGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGC 2269
          ||||||              ||||||||||||||||||||||||||||||||||||||||
Db   5175 TGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCCTGGGTGC 5234

Qy   2270 CCGCCCACAAGGGCATCGGCGGCAACGAGCAGGTCGAGCAAGAGTCGGTGCAGGGGCATCC 2329
          |||||||||||||||||              ||||||||||||||||||||||||||||||
Db   5235 CCGCCCACAAGGGCATCGGCGGCAACGAGCAGGTCGAGCAAGAGTCGGTGCAGGGGCATCC 5294

Qy   2330 GCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTA 2381
          |||||||||||||||||||||              ||||||||||||||||||
Db   5295 GCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTA 5346
```

```
RESULT 12
LOCUS       AX427931      9169 bp    DNA    linear   PAT  20-JUN-2002
DEFINITION  Sequence 169 from Patent WO0232943.
ACCESSION   AX427931
VERSION     AX427931.1  GI:21538018
KEYWORDS    
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 169 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9169
```

```
Qy   2030  ACGTGACCGACCGGGCCGGCCGCGAGAGATCGTGAGCCTGACCGAGACCAACCAACAGAAGA  2089
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   5003  ACGTGACCGACCGGGCCGGCCGCGAGAGATCGTGAGCCTGACCGAGACCAACCAACCAGAAGA  5062

Qy   2090  CCGAGCTGCAGGCCATCGAGCTGCCCTGCAGGACAGCGCAGCGGCAGCGGAGGTGAACATCGTGA  2149
           ||||||||||||||||| |||||||||||||||||||||||||||||| |||||||||||||||
Db   5063  CCGAGCTGCAGGCCATCACCTGGCCCTGCAGGACAGCGCAGCGGCCTGGAGGTGAACATCGTGA  5122

Qy   2150  CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCGTGAAGGGCAGCC  2209
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   5123  CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCGACCAGAGGCGAGC  5182

Qy   2210  TGGTGACCAGATCATCGAGCAGTCGATCAAGAAGAGAGGTGTACTTGAGCTGGGTGC  2269
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   5183  TGGTGAGCCAGATCATCGAGCAGTCGATCGATCAAGAAGAGAGGTGTACCTGGGTGC  5242

Qy   2270  CCGCCCACAAGGGCATCGGCGGCAACGAGATCGAACTGGTGAGCAAGGGCATCC  2329
           ||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   5243  CCGCCCACAAGGGCATCGGCGGCAACGAGGTGACGGGCGTGTGAGCCGCCCATCC  5302

Qy   2330  GCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTA  2381
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5303  GCAAGGTGGCTGTTCCTGGACGGCATCGACAAGGGCCCAGGAGGCACGAGAA  5354
```

RESULT 13
LOCUS   AX427926   9194 bp   DNA   linear   PAT 20-JUN-2002
DEFINITION   Sequence 164 from Patent WO0232943.
ACCESSION   AX427926
VERSION   AX427926.1 GI:2153013
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
               artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE   Modifications of hiv env, gag, and pol enhance immunogenicity for
               genetic immunization
  JOURNAL   Patent: WO 0232943-A 164 25-APR-2002;
               GOVERNMENT OF THE UNITED STATES (US)
FEATURES   Location/Qualifiers
     source   1..9194
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match   80.4%;   Score 1975;   DB 6;   Length 9194;
Best Local Similarity   90.9%;   Pred. No. 4.4e-198;
Matches 2177;   Conservative   0;   Mismatches 190;   Indels   28;   Gaps   6;

```
Qy     14  TGGCCGAGGCCATGAGCGAGGCCAGG---CCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACT    70
           ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db   2967  TGGCCGAGGCCATGAGCGAGGCCAGGTGACCAACAGCGCAGTGACCAACATGATGCAGCGCGGCAACT  3026

Qy     71  TCAAGGGCCCCAAGCGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGA   130
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   3027  TCCGCAACCAGCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACCACCCCC  3086

Qy    131  GCAACTGCCGCGCCCCCCGCGCCGCGCAGGCTGTTGCTGGAGAAGGCCAAGGAGGAGGGCAAGGAGGCCACCAGA   190
           ||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db   3087  GCAACTGCCGCGCCCCCCGCGCCGCGCAGGCTGTTGCTGGAGAAGGCCAAGGAGGAGGGCAAGGAGGCCACCAGA  3146

Qy    191  TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGAAGATCGCGGCCTTCCCCCAGG   250
           ||||||||||||| ||||||| ||||||||||| |||||||||||  ||||||||||||||||
Db   3147  TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGAAGATCGCGGCCTTCCCCACAAG  3205

Qy    251  GCAAGGCCCGCGAGTTCCCCAGCAGGACCAAGATCGGCCAAGGCCGCGCCACCAGCGCCGGAGC   310
           |||||||||||||| |||||||||||||||| ||||||||||||||||||||||||| ||||||
Db   3206  GGAAGGCCCGCCAGGGAATTTTCTTCAGAGCAGCAGACCAGCAGCCCCCCACAAGCCAACAGCGC  3265
```

```
Qy    962  CCGCCGGCCTGAAGAGAAGAAGAGAGCGTGACCGTGCTGGACGTGGCGACGCCTACTTCA  1021
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   3923  CCGCCGGCCTGAAGAGAAGAAGAGAGCGTGACCGTGCTGGACGTGGCGACGCCTACTTCA  3982

Qy   1022  GCGTGCCCCTGACGAAGTACACCGCCTTCACCATCCCCAGCATCAACA  1081
           ||||||||||||| |||||||||||||||||||||||||||||||||
Db   3983  GGGTGCCCCTGACAAGTACACCGCCTTCACCATCCCCAGCATCAACA  4042

Qy   1082  ACGAGACCCCGGCATCCGCTACCAGTACAAGACGTGTGCCCAGGCGTGAAGGGCAGCC  1141
           ||||||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db   4043  ACGAGACCCCGGCATCCGCTACCAGTACAAGACGTGCTGCCCAGGGCTGGAAGGGCAGCC  4102

Qy   1142  CCAGCATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCG  1201
           ||||||||||| ||||||||||||||| ||||||||| ||||||||||||||||||||||
Db   4103  CCGCCATCTTCAGTGAGCATGCATGAACCAAGATCCTGAGCCCTTCCGCAAGCAGAACCCCG  4162

Qy   1202  AGATCGTGATCTACCA------GGCCACCACCTGTACGTGGGCAGCGACCTGGAGATCGGGCC  1255
           |||||||||||||||||      ||||||||||||||||||||||||||||||||||||||||
Db   4163  ACATCGTGATCTACCAGTACAATGGCCACCACCTGTACGTGGGCAGCGACCTGGAGATCGGCC  4222

Qy   1256  AGCACCGCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGGCTGGGGCTTCACCACCC  1315
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   4223  AGCACCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGGCTGGGGCTTCACCACCC  4282

Qy   1316  CCGACAAGAAGCACCAGAAGAGGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCCG  1369
           |||||||||||||||||||||||||| |||||||||||||||      |||||||||||||
Db   4283  CCGACAAGAAGCACCACCAGAAGAGCCCCCCTTCCTGTGTGATGGCTACGAGCTGCACCCCG  4342

Qy   1370  ACAAGTGACCGTGCGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGACCGTGAACGACA  1429
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   4343  ACAAGTGACCGTGCGCAGCCCATCGAGCTGCTGCCCGAGAAGGAGAGCAGCTGAACGACA  4402

Qy   1430  TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGC  1489
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   4403  TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACGGCGCCGGCATCAAGGTGC  4462

Qy   1490  GCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCG  1549
           ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   4463  GCCAGCTGTGCAAGCTGCTGCGCGGCGCCACCAAGGCCCTGACCGGTGTGCCCCTGACCG  4522

Qy   1550  AGGAGGGCCAGCTGGCGAGCTGGGCCGAGAACCGCGAGATCCGTGCGGAGCCCGTGCACGGCG  1609
           |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   4523  AGGAGGGCCAGCTGGCGAGCTGGGCCGAGAACCGCGAGATCCGTGCGGAGCCCGTGCACGGCG  4582

Qy   1610  TGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCACGGCCACGACCAGT  1669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4583  TGTACTACGACCCCAGCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGCCAGGGCCCAGT  4642

Qy   1670  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGA  1729
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   4643  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCCCA  4702

Qy   1730  TGCCCGAGGCCATGAGCGAGGCCAGG---CCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACT  1789
           |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db   4703  TGCCCGAGGCCATGAGCGAGGCCAGGTGACCAACAGCGCAGTGACCAACATGATGCAGCGCGGCAACT  4762

Qy   1790  TCAAGGGCCCCAAGCGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGA  1849
           |||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   4763  TCCGCAACCAGCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACCACCCCC  4822

Qy   1850  GCAACTGCCGCGCCCCCCGCGCCGCGCAGGCTGTTGCTGGAGAAGGCCAAGGAGGAGGGCAAGGAGGCCACCAGA  1909
           ||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||
Db   4823  GCAACTGCCGCGCCCCCCGCGCCGCGCAGGCTGTTGCTGGAGAAGGCCAAGGAGGAGGGCAAGGAGGCCACCAGA  4882

Qy   1910  TGAACACCCCGGTGTGGAAGCTGGTGTGTTACCAGTGTGGTGAAGAAGGAGCCCATCATCGGCGG  1969
           ||||| ||||||||||||||||||||| ||||||| ||||||||||||||||||||||||||||
Db   4883  TGAACACCCCGGTGTGGAAGCTGGTGTGGGTTACCAGTGTGAAGAAGAAGGAGCCCATCATCGGCGG  4942

Qy   1970  CCGAGACCTTCTACGTGGACGGAGCCGCGCGCAACGCCGGCGCAACCAAGCTGGGCACGGCCGGCT  2029
           ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   4943  CCGAGACCTTCTACGTGGACGGCGACGGCGCGCAACCAAGCTGGGCACCGAAGAGAGCGCCGGCT  5002
```

```
Qy   311  TGCAGGTGCGCG------CGACAACCCCCGCAGCGAGGCGGCGGCGCCGAGCGCCAGGGCA  364
Db  3266  TTCAGGTTTGGGGAAGACAACAACTCCCTCTTCAGAGACGAGGCAGGAGCCGATAGACAAGGAA  3325

Qy   365  ------CCCTGAACTCTCCCCAGATCACCCTGGTCGCAGCGCCCTGGTCGCAGCGCATCAAGG  418
Db  3326  CTGTATCCTTTTAGCTTCCCTCAGATCATCACTCTTTGGCAGCGACCCCTGCTTCACCAANTAAAGA  3385

Qy   419  TGGGCGGCCAGATCAAGGACAGGGCCCTGCTGCGGCCGACGACCACCGTGCTGGGAGG  478
Db  3386  TAGGGGCAGCTGAAGGAGGGCCCTTTGAGACACCGGCGCCACCGTGGTGCGGCTGGAGG  3445

Qy   479  AGATGAGCGTGCCGGCAAGTGCTGCCGGGCAAGATGATCGGCGCATCGGCGCTTCATCA  538
Db  3446  AGATGAACCTGCGCGGCCCGGCCCTGGAAGCCAAGATGATCGGCGCATCGGCGGCTTCATCA  3505

Qy   539  AGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGGGCGCAAGAGGCCATCGGCCACCG  598
Db  3506  AGGTGGGGCAGTACGACCAGATCCTGATCGAGATCTGCGCCGGCCACAGAGGCCATCGGCCACCG  3565

Qy   599  TGCTGATCGGCCCCACCCCGTGAACATCGGCCGGCAACATGCTGACCCAGTCTGGGCT  658
Db  3566  TGCTGGTGGGCCCCACCCCGTGAACATCATCGGCCCGCAACCTGCTGACCCAGATCGGGCT  3625

Qy   659  GCACCCTGAACTTCCCCATCAGCCCCGTGAAGACCGGTGAAGCCCCGGCA  718
Db  3626  GCACCCTGAACTTCCCCATCAGCCCCGTGCCCGTGAAGCTGAAGCCCCGGCA  3685

Qy   719  TGGACGGGCCCAAGGTGAAGCAGTGAGCAGTGGCCCTGACCGAGAGAGATCAAGGGACCG  778
Db  3686  TGGACGGGCCCAAGGTGAAGCAGTGAGCAGTGGCCCTGACCGAGAGAGATCAAGGGCCTGTGG  3745

Qy   779  CCATCTGGGAGGACCGGCCTGAAGGAGGATGGAGGATCAGGAACCCCT  838
Db  3746  AGATCTGACCGGAGATGAGAAGGAGATGAGGAGGATCAGGAACCCCT  3805

Qy   839  ACAACACCCCGTGTTCGCCATCGAAGAAGACAGCAAGTGGCGCAAGTGGTGG  898
Db  3806  ACAACACCCCGTGTTCGCCATCGAAGAAGACAGCAAGTGGCGCAAGTGGTGG  3865

Qy   899  ACTTCCGCGAGTCGAACAAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC  958
Db  3866  ACTTCCGCGAGTCGAACAAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC  3925

Qy   959  ACCCCGCCGGCCGGCCGGCCTGAGAAGGAGAGCCGTGACCGTGCTGGACGTGGCGACGCCTACT  1018
Db  3926  ACCCCGCCGGCCGGCCGGCCTGAGAAGGAGAGCCGTGACCGTGCTGGACGTGGCGACGCCTACT  3985

Qy  1019  TCAGCGTGCCCCTGGACAGGAGACTTCCGCAAGTACACCGGCCTTCACCATCCCCAGATCA  1078
Db  3986  TCAGCGTGCCCCTGGACAGGAGACTTCCGCAAGTACACCGGCCTTCACCATCCCCAGATCA  4045

Qy  1079  ACAACGACAGCACCCCCGGCCATCGGCGTGCTGCCGCACTGACTACACAACGGCTGGAAGGGCA  1138
Db  4046  ACAACGACAGCACCCCCGGCCATCGGCGTGCTGCCGCACTGACTACACAACGGCTGGAAGGGCA  4105

Qy  1139  GCCCCAGCATCTTCAGTCTGATCATCG  1198
Db  4106  GCCCCAGCATCTTCCAGTCAGTCAGTCTGATCATCG  4165

Qy  1199  CCGAGATCGTGATCTACCA------GGCCCCCGTACGTGGGCAGGGCTGAAGGGCA  1252
Db  4166  CCGAGATCGTGATCTACCACATGAGCAGCAGGCCCCCCGTACGTGGGCAGCAGCAGGGCA  4225

Qy  1253  GCCAGCACCGCCACCGAC---CAAGATCACCACCGAAGCACCGCTGGGCGGCTTCACCA  1312
Db  4226  GCCAGCACCGCCACCGACAAGATCACCAGTCCAGCACCGTGCGCCAGCACCGCTGCGCCTTCACCA  4285

Qy  1313  CCCCCGACAAGAAGCACCGAAGAGCAGCCCCTTCCTGCCCAT------CGAGCTGCACC  1366
Db  4286  CCCCCCGACAAGAAGCACCACCGAAGAGCAGCCCCCCCTTCCTTGTGGGATGGGCTACGAGCTGCACC  4345

Qy  1367  CCGACAAGTGGACCTGGACCTGCAGCGCCATCGAGCTGCCCCGGACGTGCCCGAGAAGGAGAGTCTGACCGTGAACG  1426
```

```
Db  4346  CCGACAAGTGGACCTGCAGCGCCATCGTGCTGCCCGAGAAGGACGACGCTGGACCGTGAACG  4405
Qy  1427  ACATCCAGAAGCTGGTGTGGGCAAGCTGAACTGGGCGGCGCCAAGGCCCTGACGACATCGTGCCCCTGA  1546
Db  4406  ACATCCAGAAGCTGGTGTGGGCAAGCTGAACTGGGCGGCGCCAAGGCCCTGACCGAGGTGGTGCCCCTGA  4465

Qy  1547  CCGAGGAGGCCGAGCTGGACGTGCCGAGAACCGCGAGATCCTGCGCGAGGACCGGTGCACG  1606
Db  4466  CCGAGGAGGCCGAGCTGGACGTGCCGAGAACCGCGAGATCCTGCGCGAGGACCGGTGCACG  4525

Qy  1607  GCGTGTACTACGACCCCAGCAAGGACCTGTGGCCGGCAAGAGGCCACAGAGGCCACGACC  1666
Db  4526  GCGTGTACTACGACCCCCAGCAAGGACCTGTGGCCGGCAAGAGGCGGGCCACCAGGGGCC  4645

Qy  1667  AGTGGACCTACCAGATCTACCAGGAGCCCTTCAGGAACCTGAGACCGGCAAGTACGCCA  1726
Db  4646  AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCC  4705

Qy  1727  AGATGCGCACCGCCACCGTGATCTGGGGCAAGAGCAGTGACCCCCAAGG  1786
Db  4706  GCATGAAGGGCGCCACCGCCATCGTGGGCAAGAGCAGTGACCCCCAAGG  4825

Qy  1787  CCATGGAGACCTGGTGACCACTACTGGCAGGTCACCCTGGATCCCCGAGTGGGAGT  1906
Db  4766  CCACCGAGAGCATCGTGATCGTGACCGAGAGCAGTTCAAGCTGCCCATCCAGAAGG  4766

Qy  1847  AGACCTGGGAGACCTGGTGACCACTACTGGCAGGTCACCCTGGATCCCCGAGTGGGAGT  1906
Db  4826  AGACCTGGGAGACCTGGTGACCACTACTGGCAGGTCACCCTGGATCCCCGAGTGGGAGT  4885

Qy  1907  TCGTGAACACACCCCCCCTGTGAGCGTGTGTACCAGCTGGAGAGAGCCCCATCATCG  1966
Db  4886  TCGTGAACACACCCCCCCTGTGAGCGTGTGTACCAGCTGGAGAGAGCCCCATCATCG  4945

Qy  1967  GCGCCGAGACCTTCTACGTGGACGCGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCG  2026
Db  4946  GCGCCGAGACCTTCTACGTGGACGCGCGCCGCCAACCGCGAGACCAAGCTGGGCAAGGCCG  5005

Qy  2027  GCTACGTGACCGACCGGCCCAGAAGATCGTGAGCGTGACCGAGAAGATCGTGCCCCTGA  2086
Db  5006  GCTACGTGACCGACCGGCCCAGAAGATCGTGAGCGTGACCGAGAAGATCGTGCCCCTGA  5065

Qy  2087  AGACCGAGCTGCAGGCTGCAGCGCCATCAGCTGGCCCTGCAGGACACGCGGCAGCGAGTGAACATCG  2146
Db  5066  AGACCGAGCTGCAGGCTGCAGCGCCATCATCGGCGGGCCCTGCAGGACACGCGGCAGCGAGTGAACATCG  5125

Qy  2147  TGACCGACACGCAGCCAGTACGCCCTGGCATCATCAGCCCTGGAAGGGCA  2206
Db  5126  TGACCGACAGCAGTACGCCCTGGCATCATCAGCCCTGGAAGGGCA  5185

Qy  2207  AGCTGGTGAACCAGCAGCATCATCGGCGCCAACAGCGTGTACTGTGGCTGGCAAGGGCA  2266
Db  5186  AGCTGGTGAACCAGCAGCATCATCGGCGCCAACAGCAGTGTACTGTGGCTGGCAAGGGCA  5245

Qy  2267  TGCCCGCCCAGAAGGCATCGATCGTGTCCTGGCAAGGCGACCTTGGCGGCATCGGCGATCG  2326
Db  5246  TGCCCGCCCACAAGGCATCGATCGTGTCCTGGCAAGGCGACCTGGGCGGCATCGGCGATCG  5305

Qy  2327  TCCGCAAGGTGCTGTTCCTGACGCTGCTGCTGGGCGATCGCGTGTGTCGTGTAC  2381
Db  5306  TCCGCAAGGTGCTGTTCCTGACGCTGCTGCTGGGCGATCGCGTGTGTCGTGTAC  5360
```

```
VERSION     AX427925.1  GI:21538012
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 163 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES            Location/Qualifiers
     source         1..9194
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
                    /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match             80.3%;  Score 1973.4;  DB 6;  Length 9194;
Best Local Similarity   90.9%;  Pred. No. 6.5e-198;
Matches 2176;  Conservative  0;  Mismatches 191;  Indels 28;  Gaps  6;
```

```
Qy      14  TGGCCGAGGCCATGAGCAGGCAGG---CCACCAGCGCGCAACATCCTGATGCAGCGCAACT   70
                ||||||||||||||||||     ||||| || ||| | |||||||||||  ||||||
Db    2967  TGGCCGAGGCCATGAGCCAGGTGACCAACAGCAGCGCCACCATCATGATGCAGCGCGGCAACT  3026

Qy      71  TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCCACATCGCCC  130
                ||||||||||||||||||  ||||||||||||||||| |||||||||||||| ||||
Db    3027  TCCGCAACCAGCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACACCCGCCC  3086

Qy     131  GCAACTGCCGGCCGCCCCCGCGGAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAGGGCCACCAGA  190
                ||||||||||  ||||||||||      | |||| |||||| | ||||||||||| ||||||
Db    3087  GCAACTGCCGCGCGCGCCCGCGGAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAGGGCCACCAGA  3146

Qy     191  TGAAGGACTGCACCGAGCGCAGCCGCCAGCCCAACTTCTTCGCGGAGGACCTGGCCTTCCCCCAGG  250
                ||| ||||||||||||||| ||||||||||| ||||||||||||  |||| ||||| |||||||
Db    3147  TGAAGGACTGCACCGAGCGCAGCAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCACCACACAAG  3205

Qy     251  GCAAGGCCGCGCGAGTTCCCCAGCGACAGCAGAACCGCGCAACAGCCCACCAGCGCCGCAGAGC  310
                | ||||||||||||||       |||| |||||||   ||||| |||||||| ||||||||
Db    3206  GGAAGGCCCAGGGAATTTTCTTCAGGGCAGGACCAGAGCCAACAGCCCCACCACCAGAAGAGAGC  3265

Qy     311  TGCAGTGCCGCG-----CGACAACCCCCGGCAGCCCGCGGAGGCGCGGCCCAGGGCA  364
                ||||| |||||     |||| |||   |||||| ||| |||| |||| |||||||
Db    3266  TTCAGGTTTGGGGAAGAGACAACAACACTCCCTCTCAGAAGCAGGAGCCGATAGACAAGGAA  3325

Qy     365  -----CCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCCCCCTGGTTGAGCATCAAGG  418
                     | |||| || ||||||||||||||||||||||||   ||||||| ||||||||
Db    3326  CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGA  3385

Qy     419  TGGGCGGCCAGATCAAGGAGGCCCTGCTGGAACCGGCGCCGACACACCGTGCGGCGGCATCA  478
                ||||| |||||| | ||||| ||||| ||||| |||| || |||||||||  |||||| |||
Db    3386  TAGGGGGCCAGCTGAAGGAGGCCCTGTTAGACACCGGCGCACACGACACCGTGCTGGAGGA  3445

Qy     479  AGATGAGCCTGCCCGGGCGCTGGAAGCCCGCTGGAGTCCGGCCCATCGCGCGGCGGCTTCATCA  538
                |||||| ||||||||||||||  |||| ||||||||| || ||||||  ||||||| |||||
Db    3446  AGATGAACCTGCCAGGAAGATGGAAGCCCAAAATGATCGGCGGAATCGGCGGTTCCATCA  3505

Qy     539  AGGTGCGGCCAGTACGACCAGATCTGAGCGATGGAGATCGAGAGGCCCATCGGCACCG  598
                ||||| |||||||||||||  ||||||||| ||||||||||| ||||||||||| ||||
Db    3506  AGGTGGGGCCAGTACGACCAGATCTGAGCGATCATCGGCGCCACAGGGCCATCGGCACCG  3565

Qy     599  TGCTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGTGGGGCT  658
                || ||||||| ||||||||||||||  || ||||| |||||||||||||||| |||||||
Db    3566  TGCTGGTGGGCCCCACCCCCGTGAACATCATCGGCCGCAACCTGCTGACCCAGATCGGCT  3625

Qy     659  GCACCCTGAACTTCCCCCATCAGCCCCATCCGCCCCTGCCCGTGAAGTGAAGCCCGGCA  718
                ||||||||||| ||||||  || |||||| | |||| ||||||||| |||||||||||||
Db    3626  GCACCCTGAACTTCCCCATCAGCCCCATCCGCGACCTGCCCGTGAAGTGAAGCCCGGCA  3685

Qy     719  TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG  778
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3686  TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGG  3745
```

```
Db    3686  TGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGG  3745
Qy     779  CCATCTGCGAGGATGGAGAAGGAGGGCAAGATCCAAGATCCGGCCCGAGAACCCCT  838
Db    3746  AGATCTGCACCGAGGATTGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCGAGAACCCCT  3805
Qy     839  ACAACACCCCGTTCGCCATCAAGAGAGAGGAGAGCACCAAGTGCGCAAGCTGGTGTG  898
Db    3806  ACAACACCCCGTTCGCCATCAAGAGAGAGGAGAGCACCAAGTGGCGCAAGCTGGTGTG  3865
Qy     899  ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC  958
Db    3866  ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC  3925
Qy     959  ACCCCGCCGGCCTGAAGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCGCTGCCCGCTACT  1018
Db    3926  ACCCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCGCCTACT  3985
Qy    1019  TCAGGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCGTCCCAGCATCA  1078
Db    3986  TCAGGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA  4045
Qy    1079  ACAACGAGACCCCCGGCATCCGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA  1138
Db    4046  ACAACGAGACCCCCGGCATCCGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA  4105
Qy    1139  GCCCCCAGATCTTCCAGAGCAGGATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACC  1198
Db    4106  GCCCCGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACC  4165
Qy    1199  --------GGCCCCCCGTACGTGGGCAGCGACACTGGGAGATCG  1252
Db    4166  CCGACATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCG  4225
Qy    1253  GCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCCTTCACCA  1312
Db    4226  GCCAGCACCGCGCCAAGATCGAGGAGCTGCGCCAGCACCTGCTGCGCTGGGGCCTTCACCA  4285
Qy    1313  CCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGCCCAT------CGAGTGCGACC  1366
Db    4286  CCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGATGGGGATGGAGAGAGTGGAACG  4345
Qy    1367  CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGTGGAACG  1426
Db    4346  CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACG  4405
Qy    1427  ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGG  1486
Db    4406  ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGG  4465
Qy    1487  TGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGA  1546
Db    4466  TGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGAGGTGGTGCCCCTGA  4525
Qy    1547  CCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGTGCACCG  1606
Db    4526  CCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGTGCACCG  4585
Qy    1607  GCGTGTACTACGACCCAGAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGA  1666
Db    4586  GCGTGTACTACGACCCAGAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGC  4645
Qy    1667  AGTGGACCTACGGCGCCGGCTGAAGCTGCGGGCCGCCGCCCAGCTACGCCCGAGAGATCG  1726
Db    4646  AGTGGACCTACGGCGCCGGCTGAAGCTGCGCGGCCGCCGCCCAGCTACGCCGAGAGATACG  4705
Qy    1727  AGATGCGCACCGGCCGCCCACCACCAACACCAACCCAAGTTCGGGCTGCCTGCCCATCCAGAAGG  1786
Db    4706  GCATGAAGGGCGGCCCACCACCAACACCAACACCAAGTTCGGGCTGCCTGCCCATCCAGAAGG  4765
Qy    1787  CCATCGAGAGCATCGTGATCTGGGGCAAGAAGATCAAGGCCCTGAGTTCCGCCTGCCCATCCAGAGG  1846
Db    4766  CCACCGAGAGCATCGTGATCTGGGGCAAGAAGATCAAGGCCCTGAGTTCAAGTTCAAGCTGCCATCCAGAGG  4825
```

Qy   131 GCAACTGCCGCGCCCCGCAAGAAGGGTCTGTGGAACTGCTGGAGTGCGGCAAGGAGGGGCCACCAGA 190
          |||||||||||| |||||||||||| |||| ||||| ||||||||||||||||||| ||||||||||
Db   3087 GCAACTGCCGCGGCCCCGCAAGAGGGTGCTGGAAGTGCGGCAAGGAGGAGGGCCACCAGA 3146

Qy   191 TGAAGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGGGAGAAGAGGAGCCCATCATCG 250
          |||||||||||||||||||||||||| |||| ||| || ||||| |||||||
Db   3147 TGAAGGACTGCACCGAGCGGCCAGGGACAGGCTAA-TTTTTAGGGAAGAATCTGCCTTCCCACAAG 3205

Qy   251 GCGCCCGAGACCTTCTACGTGACGGACGCGGCCAACCGCGAGACCAAGATCGGCAAGGAGCC 310
          || ||||||| | |||| |||| |||||| |||||| ||| |||||||||||||||| |||||
Db   3206 GGAAGGCCAGGGAATTTCTTCAGGACGCACAGCCAACAGCCCCACCAGAGAGAGC 3265

Qy   311 GCTACGTGACCGACGGGCGCCGAGAGATCGTGAGCCTGACCGACCAGACCAGA 364
              |||||||||||| ||||| ||||| ||| || | ||||| |||||
Db   3266 TGCAGTGCGCGG-----CGACAACCCCGCA|||||TCCCTC TCAGAGCAGGCAACAACTCCCTCTCAGAGCAGGCCGATAGACAAGAA 3325

Qy   365 -----CCCTGAACTTCCCCGAATCACCCTGTGCAGCCAGGTTGAACATCG 418
                   |||||| ||| | | | | || ||||||
Db   3326 CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACACCGAGGGTTGAACATAAGA 3385

Qy   419 TGGGCGGCAGATCAAGGAGCCCGCTGCTGGACACCGGCGCGACACCGTGCTGGAGG 478
          ||||| |||||| |||||| ||| || |||||| |||||||| |||| |||||||||
Db   3386 TAGGGGGCCAGCTGAGGAGCCCTTTCAGACACCGGCCCAGCCCCGACACCGTGCTGGAGG 3445

Qy   479 AGATGAGCCTGCCCGGCAGATCAAGCACCGGCCAAGAGGCCATCGGCGGCTTCATCA 538
          ||||||||||| |||||| ||| | ||| ||||| |||||||| ||||||||||
Db   3446 AGATGAGCCTGCCCGGCAGATCCTGATCTGCGGCTCCGGCCCAAGATGATCGGCGGCTTCATCA 3505

Qy   539 AGGTCGCCAGTACGACCAGATCTGATCGAGATCCGGCCGCAAGAAGGCCATCGGCACCG 598
          |||| ||||||| || ||||||||| ||||||||||||||||||||||| |||||||||
Db   3506 AGGTGGGCCAGTACCAGATCTGATCTGATCCGGCCGCAACAAGGCCATCGGCACCG 3565

Qy   599 TGCTGATCGGCCCCACCCCCGTGAACATCAGCCCATCGAGACCGGTGCCGTGAAGCTGAAGCCCGGCT 658
          ||||||| ||||||||||||||||||||||| |||||| |||| ||| ||| | |||||||||| |||
Db   3566 TGCTGGTGGGCCCCACCCCCGTGAACATCATCATCGGCCCGAACAATCGGTGCCGTGAAGCTGAAGCCCGGCT 3625

Qy   659 GCACCCTGAACTTCCCCATCAGCCCATCGAGACCGGTGCCGTGAAGCTGAAGCCCGGCA 718
          |||||||||||||||||| ||||| | |||| | |||||| |||| |||||||||||||
Db   3626 GCACCCTGAACTTCCCCATCAGCCCATCGAGACCGGTGCCGTGAAGCTGAAGCCCGGCA 3685

Qy   719 TGGACGGGCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGATCAAGGCCCTGACCG 778
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   3686 TGGACGGGCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGATCAAGGCCCTGTGG 3745

Qy   779 CCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCT 838
          ||| |||||||||||| |||||||||||||||||| |||||||||||| |||||||| |||
Db   3746 AGATCTGCACCGACGATGGAGAGGAGGGCCAAGATCAGCAGGATCAGGCGCCGCGAGAACCCCT 3805

Qy   839 ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
          ||||||||||| ||||| |||||| ||||| |||||||||||||||||| |||||||||
Db   3806 ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG 3865

Qy   899 ACTTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 958
          ||||||| |||||||||||||| ||||||||||||||| ||||||||||||||||||||||
Db   3866 ACTTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 3925

Qy   959 ACCCCGCCGGGCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGGCACGCCTACT 1018
          |||||| ||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db   3926 ACCCCGCCGGGCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGGCACGCCTACT 3985

Qy   1019 TCAGCGTGCCCCTGACGACGAGGACTTCCGGAAGTACACCGCCTTCACCATCCCCAGCATCA 1078
          ||||||||||||||||||| ||||||||||||| |||||||||||||| ||||||||||||||
Db   3986 TCAGCGTGCCCCTGACGACTTCCGGAAGTACACCGCCTTCACCATCCCCAGCATCA 4045

Qy   1079 ACAACGAGCACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA 1138
          |||||||| |||||||||||||||||||||||| || || |||||||||||| || ||||
Db   4046 ACAACGAGACACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA 4105

Qy   1139 GCCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACC 1198
          || ||||||||||||||| ||||| |||| || ||| ||||||||||| |||||||
Db   4106 GCCCCGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAAGCACAGAACC 4165

Qy   1847 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGATCCCCGAGTGGGGAGT 1906
          |||||||||||||||||||||||| ||||||||||||||| |||||||||||||||||||
Db   4826 AGACCTGGGAGGCCTGGTGGACCGACTACTGGCAGGCCACCTGATCCCCGAGTGGGGAGT 4885

Qy   1907 TCGTGAACACCCCCGCCCCTGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCG 1966
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   4886 TCGTGAACACCCCCGCCCCTGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCG 4945

Qy   1967 GCGCCCGAGACCTTCTACGTGACGGACGCGGCCAACCGCGAGACCAAGATCGGCAAGGCCG 2026
          || || ||||||| |||||||| |||||||||||||| |||||||||||||||| ||||||
Db   4946 GCGCCCGAGACCTTCTACGTGACGGACGCGGCCAACCGCGAGACCAAGATCGGCAAGGCCG 5005

Qy   2027 GCTACGTGACCGACGGGCGCCGAGAGATCGTGAGCCTGACCGACCAGACCAGA 2086
          |||||||||||||| |||||||| |||| || ||||||||||||||||||||||
Db   5006 GCTACGTGACCGACGGGCCGCCAGGAAGGTGGTGCCCTGACCGACCACCAACCAGA 5065

Qy   2087 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCTGGAAGTTGAACATCG 2146
          ||||| ||| |||||| |||||| ||||| ||||||| || || || |||||||||||||||
Db   5066 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACGCGGCCTGGAGGTGAACATCG 5125

Qy   2147 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCG 2206
          ||||||||||||||| |||||||||||||||||||| |||||||||||||||||||||| |
Db   5126 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCG 5185

Qy   2207 AGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGG 2266
          |||||||||| |||||||||||||||||| |||||||| ||| ||||||||||||||||
Db   5186 AGCTGGTGGAGCAGATCATCGGCCGGCATCTGATCAAGGAGAAGGTGTACCTGGCCTGGG 5245

Qy   2267 TGCCCCGCCACAAGGGCATCGGCGGCAACGAGATCGACAAGCTGGTGAGCGAAGGGCA 2326
          ||||| ||||||||||||||||||| |||||||||| |||||| || |||||| |||||
Db   5246 TGCCCGCCCACAAGGGCATCGGCGGCAACGAGTGGACAGTGGACGCGCCTGGTGAGCGGGCCGCA 5305

Qy   2327 TCCGCAAGGTGCTGTTCCTGGACGGCATCGATGGGCATCGTGATCTACCAGTA 2381
          | ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   5306 TCCGCAAGGTGCTGTTCCTGGACGGCATCGATGGGCATCGTGATCTACCAGTAAAA 5360

RESULT 15
AX427927
LOCUS   AX427927   12411 bp   DNA   linear   PAT 20-JUN-2002
DEFINITION Sequence 165 from Patent WO0232943.
ACCESSION   AX427927
VERSION   AX427927.1 GI:21538014
KEYWORDS   .
SOURCE   synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS   Huang,Y. and Nabel,G.J.
TITLE   Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
JOURNAL   Patent: WO 0232943-A 165 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES   Location/Qualifiers
source   1..12411
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match   80.3%; Score 1973.4; DB 6; Length 12411;
Best Local Similarity   90.9%; Pred. No. 6e-198;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps   6;

Qy   14 TGGCCGAGGCATGAGCCAGG---CCACCAGCGGCCAACATCCGATGCGGCCAGCAACT 70
          |||||| ||||||||||||| |||||||||||| ||| |||| ||||||| |||
Db   2967 TGGCCCGAGGCCATGAGCCAGGTCGTGAGCCAGGCCAACAGCGCCACCATCATGATGCGCGGCCAACT 3026

Qy   71 TCAAGGGCCCCAAGCGACCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCC 130
          || | ||| |||||||| |||||||| |||||||||||||||||||||||| ||||||| ||||
Db   3027 TCCGCAACCAGCGCCAGATGTGAAGTGCTTCAACTGCGGCAAGGAGGAGGGCCACACCGCCC 3086

```
Db  5246 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
         TGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGGTGGACGGCCTGGTTGAGCGCCGGCA 5305

Qy  2327 TCCGCAAGGTGCTGTTCCTGGACGGCATCGATGCGGCATCGTGTGATCGTGATCTACCAGTA 2381
         |||||||||||||||||||||||||||   ||||||
Db  5306 TCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGAGGAGCACGAGAA 5360
```

```
Qy  1199 CCGAGATCGTGATCTACCA------GGCCCCCTGTACGTGGCGAGCGACCTGGAGATCG 1252
         |||| ||||||||||||||      |||||||||||||||||||||||||||||||||
Db  4166 CCGACATCGTGATCTACCAGTACCAGTACATGGACGACCTGTACGTGGCGAGCGACCTGGAGATCG 4225

Qy  1253 GCCAGCACCGGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCGTGGGGCTTCACCA 1312
         |||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db  4226 GCCAGCACCGGCGCCAAGATCGAGGAGCTGCGCCAGCACCTGCTGCGCGTGGGGCTTCACCA 4285

Qy  1313 CCCCCGCAAGAAGCACCAGAAGGAGCACCCCCTTCCTTCCTTGCCCAT------CGAGCTGCACC 1366
         |||| |||||||||||||||||||||||||||||||||||||||||||      ||||||||||
Db  4286 CCCCCGACAAGAAGCACCAGAAGGAGCACCCCCTTCCTTGTGGATGGGCTACGAGCTGCGCACC 4345

Qy  1367 CCGACAAGTGGACCGTGCGAGCCTGCATCGAGCTGCCCGACGCTGCCCGCCATCGCCCGTGAACG 1426
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4346 CCGACAAGTGGACCGTGCGAGCCTGCCATCGCTGCTGCCCGAGAAGGAGCAGTCAGCCGTGAACG 4405

Qy  1427 ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGG 1486
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4406 ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACGCCGGCATCAAGG 4465

Qy  1487 TGGCGCCAGCTGTGCAAGCTGCTGCGCGGCGAACCGGCAAGGCCCAGCACGTGCCCCTGA 1546
         |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4466 TGGCCCAGCTGTGCAAGCTGCTGCGCGGCACCGGCAAGGCCCGAGGTGGTGCCCCCTGA 4525

Qy  1547 CCGAGGAGGCCGAGCTGGAGCTGGACTGGACCCCGAGATCTGCGCGGAGCCCGTGCACG 1606
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4526 CCGAGGAGGCCGAGCTGGAGCTGGACTGGCCCGAGATCCTGGAAGGAGCCCGTGCACG 4585

Qy  1607 GCGTGTACTACGACCCCAGCAAGGACCTGGTGCCCGAGATCCAGAAGCAGGGCCACGACC 1666
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db  4586 GCGTGTACTACGACCCCAGCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGGCCAGGGCC 4645

Qy  1667 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGACCAGTACGCCA 1726
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4646 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACCGCCC 4705

Qy  1727 AGATGCCGCACGCGCCCACACCAACGACGTGAAGCAGTCGAAGCAGTTCCCGCTGCGAAGATCG 1786
         ||||                ||||||||||||||||||||||||||||||||||||||||||
Db  4706 GCATGAAGGGCGCCCACACCAACGACGTGAAGCAGTCGAAGCAGTTCCCGAGTGGGAGT 4765

Qy  1787 CCATGGAGAGCAGCATCGTGATCGTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGG 1846
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4766 CCACCGAGAGCAGCATCGTGATCGTGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAGG 4825

Qy  1847 AGACCTGGGAGACCACGCGCCCACCAACGACGTGAAGCAGTCGAAGCACGACCGGCTGCGCGAGATCG 1906
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4826 AGACCTGGGAGACCACGCGCCCACCAACGACGTGAAGCAGTCGAAGCACGACCGGCTGGGGAGT 4885

Qy  1907 TCGTGAACACCCCCCCCCCGTGAAGCTGTGGTACCAGCTGTGGAGAAGGAGCCCATCATCG 1966
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4886 TCGTGAACACCCCCCCCCCGTGAAGCTGTGGTACCAGCTGGGGAAGGAGCCCATCATCG 4945

Qy  1967 GCGCCGAGACCTTCTCACGTGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCG 2026
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4946 GCGCCGAGACCTTCTCACGTGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCG 5005

Qy  2027 GCTACGTGACCGGCGCGGGCGCCGGCAGAAGATCGTGAGCCTGACCGGAGACACCAACCAGA 2086
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5006 GCTACGTGACCGGCGCGGGCGCCGGCAGAAGGTGGTGCCCCTGACCGGACCACCAACCAGA 5065

Qy  2087 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGACGGCGAGGTGAACATCG 2146
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5066 AGACCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCG 5125

Qy  2147 TGACCGACAGCCAGTACGCCCCTGGGCATCATCCAGGCCCAGGATCATCCAGGCCCGACAAGAGCGAGGGCG 2206
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5126 TGACCGACAGCCAGTACGCCCCTGGGCATCATCCAGCCCCGACAAGAGCGAGGAGCG 5185

Qy  2207 AGCTGGTGAACCAGATCATCGAGCAGTCGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG 2266
         ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db  5186 AGCTGGTGAGCCAGATCATCGAGCAGTCGATCAAGAAGGAGAAGGTGTACCTGGCCTGGG 5245

Qy  2267 TGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGGTGAGCAAGGGCA 2326
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        April 10, 2004, 12:49:38 ; Search time 591.888 Seconds
                                (without alignments)
                                15572.780 Million cell updates/sec

Title:         US-09-610-313-32
Perfect score: 2457
Sequence:      1 gtcgacgcaccatggccga.........gggctagcaccggtgaattc 2457

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      2475585 seqs, 187573O760 residues

Total number of hits satisfying chosen parameters:     4951170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     Published_Applications_NA:*
               1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
               2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
               3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
               4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
               5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
               6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
               7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
               8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
               9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
              10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
              11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
              12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
              13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
              14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
              15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
              16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
              17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
              18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	DB	Length	ID	Description
1	2443.4	99.4	14	2445	US-10-190-435-43	Sequence 43, Appl
2	2443.4	99.4	14	2445	US-10-190-305A-37	Sequence 37, Appl
3	2436.4	99.2	14	3930	US-10-190-435-9	Sequence 9, Appli
4	2434.8	99.1	14	3930	US-10-190-435-10	Sequence 10, Appl
5	2434.8	99.1	14	3930	US-10-190-435-11	Sequence 11, Appl
6	2434.8	99.1	14	5184	US-10-190-435-58	Sequence 58, Appl
7	2434.8	99.1	14	5184	US-10-190-305A-82	Sequence 82, Appl
8	2428.6	98.8	14	2457	US-10-190-435-44	Sequence 44, Appl
9	2428.6	98.8	14	2457	US-10-190-305A-38	Sequence 38, Appl
10	2404.4	97.9	14	3531	US-10-190-435-13	Sequence 13, Appl
11	2403.4	97.8	14	2457	US-10-190-435-45	Sequence 45, Appl
12	2403.4	97.8	14	2457	US-10-190-305A-39	Sequence 39, Appl
13	2402.8	97.8	14	3537	US-10-190-435-14	Sequence 14, Appl
14	2402.8	97.8	14	3537	US-10-190-435-15	Sequence 15, Appl
15	2401.8	97.8	14	5145	US-10-190-435-12	Sequence 12, Appl
16	2401.8	97.8	14	5145	US-10-190-305A-12	Sequence 12, Appl
17	2343.4	95.4	14	3607	US-10-190-435-48	Sequence 48, Appl
18	2343.4	95.4	14	3607	US-10-190-305A-42	Sequence 42, Appl
19	2325.2	94.6	14	3597	US-10-190-435-46	Sequence 46, Appl
20	2325.2	94.6	14	3597	US-10-190-305A-40	Sequence 40, Appl
21	2295.8	93.4	14	3624	US-10-190-435-47	Sequence 47, Appl
22	2295.8	93.4	14	3624	US-10-190-305A-41	Sequence 41, Appl
23	2136	86.9	14	2460	US-10-241-009-30	Sequence 30, Appl
24	2136	86.9	14	2460	US-10-190-434B-30	Sequence 30, Appl
25	2136	86.9	14	2460	US-10-190-305A-32	Sequence 32, Appl
26	2135	86.9	14	3564	US-10-241-009-13	Sequence 13, Appl
27	2135	86.9	14	3564	US-10-241-009-14	Sequence 14, Appl
28	2135	86.9	14	3564	US-10-190-434B-13	Sequence 13, Appl
29	2135	86.9	14	3564	US-10-190-434B-14	Sequence 14, Appl
30	2134.4	86.9	14	4716	US-10-190-435-17	Sequence 17, Appl
31	2134.4	86.9	14	4716	US-10-190-305A-13	Sequence 13, Appl
32	2130.6	86.7	14	3999	US-10-241-009-9	Sequence 9, Appli
33	2130.6	86.7	14	3999	US-10-190-434B-9	Sequence 9, Appli
34	2129	86.7	14	3999	US-10-241-009-10	Sequence 10, Appl
35	2129	86.7	14	3999	US-10-241-009-11	Sequence 11, Appl
36	2129	86.7	14	3999	US-10-190-434B-10	Sequence 10, Appl
37	2129	86.7	14	3999	US-10-190-434B-11	Sequence 11, Appl
38	2129	86.7	14	5283	US-10-241-009-54	Sequence 54, Appl
39	2129	86.7	14	5283	US-10-190-434B-54	Sequence 54, Appl
40	2129	86.6	14	5283	US-10-190-305A-75	Sequence 75, Appl
41	2128.8	86.6	14	4713	US-10-190-435-59	Sequence 59, Appl
42	2128.8	86.6	14	4713	US-10-190-305A-83	Sequence 83, Appl
43	2128.4	86.6	14	3462	US-10-190-435-16	Sequence 16, Appl
44	2115.2	86.1	14	2466	US-10-241-009-31	Sequence 31, Appl
45	2115.2	86.1	14	2466	US-10-190-434B-31	Sequence 31, Appl

ALIGNMENTS

```
RESULT 1
US-10-190-435-43
; Sequence 43, Application US/10190435
; Publication No. US2003014324BA1
; GENERAL INFORMATION:
; APPLICANT:    ZUR MEGEDE, Jan
; APPLICANT:    BARNETT, Susan W.
; APPLICANT:    LIAN, Ying
; APPLICANT:    ENGELBRECHT, Susan
; APPLICANT:    VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP8133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE:  2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 43
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Pol.opt.YMWM_C
US-10-190-435-43

  Query Match             99.4%;  Score 2443.4;  DB 14;  Length 2445;
  Best Local Similarity   99.4%;  Score 2443.4;  DB 14;  Length 2445;
  Matches 2444;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    7 GCCACCATGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGAGC 66
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GCCACCATGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGAGC 60

Qy   67 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 126
         |||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 120

Qy  127 GCCCGCAACTGCCGCGCCCCCGCGCCCCCCGGAAGAAGAGGCTGCTCGGAAGTGCGGCAAGAAGGAGGGGCAC 186
```

Right column:

```
1201  GTGATCTACCAGGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCAGCACCGCGCC  1260
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    AAGATCGAGGAGCTGCGCAAGACCTGCGCAGTGCGGCAAGAGAAG  1326
      ||||||||||||||||||||||||||||||||||||||||||||||
Db    AAGATCGAGGAGCTGCGCAAGCACTTCACCACCCCGACAAGAAG  1320
1267
1261

1327  CACCAGAAGCGAGCCCCCTTCTGCCATCGAGCTGCACCCGACAAGTGACCGTGCAG  1386
Db    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    CACCAGAAGGAGCCCCCTTCTGCCCATCGAGCTGCACCCGACAAGTGGACCGTGCAG  1380
1321

1387  CCCATCGAGCTGCCGACAACCCGGAGAGAGTGGACCGTGAACGACATCCAGAAGTGGTCGGGC  1446
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    CCCATCGAGCTGCCCGAGAAGGAGAGAGTGGACCGTGACCGTGAACGACATCCAGAAGTGGTGTGGGC  1440
1381

1447  AAGCTGAACTGGGCCAGCCAGATCTACCCGGCCATCAAGGTGCGCGAGCTGGAG  1506
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    AAGCTGAACTGGGCCAGCCAGATCTACCCGGCCATCAAGGTGCGCGAGCTGGAG  1500
1441

1507  CTGCGCGGAACCGGCGCAAGGCCCTGACCGACATCGTGCCCGAGCCCTGACGAGCTGAG  1566
Db    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    CTGCGCGGCGGCGCCAAGGCCCTGACCGACATCGTGCCCGAGCCCTGACGAGCTGGAG  1560
1501

1567  CTGGCCGAGAACCGCGAGATCCTGCGCGAGCTTCATCAAGGTGCGC  1626
Db    ||||||||||||||||||||||||||||||||||||||||||||||
Qy    CTGGCCGAGAACCGGCGGGTGTACTACGACCCCCAGC  1620
1561

1627  AAGGACCTGCGTGGCCGAGATCCAGAAGCAGGGGCCACGACCAGATCTACCCACCAGATCTAC  1686
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    AAGGACCTGCGTGGCCGAGATCCAGAAGCAGGGGCCACGACCAGTTGGACCTTACCCACCAGATCTAC  1680
1621

1687  CAGGAGCCCTTCAAGAACCTGAAGACCGACCAGATGCCCAAGATGGCCAAGACACC  1746
Db    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    CAGGAGCCCTTCAAGAACCTGAAGACCGACCAGATGCCCAAGATGGCCAAGACACC  1740
1681

1747  AACGACGTGAAGCAGTCAGCCGTGCCGTGCGAGAGATCGCCATGAGAGCATCGTGATC  1806
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    AACGACGTGAAGCAGTCAGCCGTGCCGTGCGAGAGATCGCCATGAGAGCATCGTGATC  1800
1741

1807  TGGGCAAGACCCCCAAGTTCCGGCTGCCCGAGTGGGAGACCTGGTGG  1866
Db    |||||||||||||||||||||||||||||||||||||||||||||||
Qy    TGGGCAAGACCCCCAAGTTCCGGCTGCCCGAGTGGGAGACCTGGTGG  1860
1801

1867  ACCGACTACTGGCAGGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCCCTG  1926
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    ACCGACTACTGGCAGGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCCCTG  1920
1861

1927  GTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCAAGATCGGCAGCCGCGGCTTCTACGTG  1986
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    GTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCAAGATCGGCAGCCGCGGCTTCTACGTG  1980
1921

1987  GACGGCGCCCAACAGATCGGCAGGACCAAGACCAGAGCTGCAGGCCATC  2046
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    GACGGCGCCCAACAGATCGGCAGCAGCGGCAGGACCAAGACCAGAGCTGCAGGCCATC  2040
1981

2047  CGGCAGAAGATCGTGAGCCTGAGCCTGCTCGAGGACACCAGAGAGCCAGGCGTG  2106
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    CGGCAGAAGATCATCCAGGCCCTGGAGCTGACCGACAGCCAGCAGGGCCATC  2100
2041

2107  CAGCTGGCCCTGCGCAGGGCAGCAGCCAGCCCAGCCGCCTCAGGGTGACCGACTACGCGC  2166
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    CAGCTGGCCCTGCGCAGGGCAGCAGCCAGCCCAGCCGCCTCAGGGTGACCGACTACGCGC  2160
2101

2167  CTGGGCATCATCCAGGACCTGAGCGGGCTGGTGGAACCAGCCAGCAGATCATC  2226
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    CTGGGCATCATCCAGGACCTGAGCGGGCTGGTGGAACCAGCCAGCAGATCATC  2220
2161

2227  GAGCAGCTGATCAAGAAGAGAAGGTGTACCTGAGCTGGCCCCACAAGGGCATC  2286
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    GAGCAGCTGATCAAGAAGAGAAGGTGTACCTGAGCTGGCCCCACAAGGGCATC  2280
2221

2287  GGCGGCAACGACAGCAGATCGACAAGTCGGTGAGCAAGGCTCGTTCCTG  2346
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    GGCGGCAACGACAGCAGATCGACAAGTCGGTGAGCAAGGCTCGTTCCTG  2340
2281
```

Left column:

```
Db    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
121   GCCCGCAACTTCCGCGCCCCCGAAGAGGGCTGCTGGAAGTGCGGCAAGAGGGGCCAC  180
Qy    CAGATGAAGGACTGCACCGAGCGCGCCAGCCAACTTCTTCCGCGGAGGACTTGGCCTTCCC  246
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CAGATGAAGGACTGCACCGAGCGCGCCAGCCAACTTCTTCCGCGGAGGACTTGGCCTTCCC  240
187
181

247   CAGGGGCAAGGCCCGGCGACGTCCCCGCAGCAGAACCGGCGCCACCAGCCCACCAGCCGC  306
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CAGGGGCAAGGCCCGGCGACGTCCCCAGCGAGCAACCGGCGCCACCAGCCCACCAGCCGC  300
241

307   GAGCTGCAGGTGCGGCGGCGACAACCCCCGCAGCGAGGCCGGCCAGCGACCACC  366
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GAGCTGCAGGTGCGGCGGCGACAACCCCGCAGCGAGGCCGGCCAGCGACCACC  360
301

367   CTGAACTTCCCCAGATCACCCTGTGCAGCGCCCCCTGTGAGCATCAAGGTGGGCGGC  426
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CTGAACTTCCCCAGATCACCCTGTGCAGCATCAAGGTGGGCATCAAGGTGGGCGGC  420
361

427   CAGATCAAGGAGGCCCTGCTGACACACCGGCCTGGAAGCCCAAGATCGCGCGGCTTCATCAAGGTGCGC  486
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CAGATCAAGGAGGCCCTGCTGACACGTCGGCCTGGAAGCCCAAGATCGCGCGGCTTCATCAAGGTGCGC  480
421

487   CTGCCCCGCAAGTGGAAGCCCAAGATCGGCGGCAAGAAGGCCATCGGCGCGGTTCATCAAGGTGCGC  546
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CTGCCCCGCAAGTGGAAGCTGAAGATCGGCGGCAAGAAGGCCATCGGCGCGGTTCATCAAGGTGCGC  540
481

547   CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC  606
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC  600
541

607   GGGCCCCCCCGTGAACATCATCTCGGCCGCCCGTGCCCGTGAAGCTGAGCCCGGCATGGACGGC  666
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GGCCCCACCCCGTGAACATCATCTCGGCCGCCCGTGCCCGTGAAGCTGAGCCCGGCATGGACGGC  660
601

667   AACTTCCCCATCAGCCCATCGAGACCCGTGCCCGTGAAGCTGAACGCCGGCCATCTGC  726
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    AACTTCCCCATCAGCCCATCGAGACCCGTGCCCGTGAAGCTGAACGCCGGCCATCTGC  720
661

727   CCCAAGTGAAGCAGTGGCCCCTGACCGAGACCAGCAGCAGTCAACGGCCCCATCGTGC  786
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CCCAAGTGAAGCAGTGGCCCCTGACCGAGACGAGCAGTCAACGGCCCCATCGTGC  780
721

787   GAGGAGATGGAGAGAGGGCCAAGATCACCAAGATGCGGCCCCCGAGAACCCCTACACACC  846
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GAGGAGATGGAGAGAGGGCCAAGATCGGCCAAGATGCGGCCCCCGAGAACCCCTACACACC  840
781

847   CCCGTGTTGCCATCAAGGACTTCTGCTGGAGGTGGCAGCTGGCCATCCCCACCCGCGC  906
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CCCGTGTTGCCATCAAGGACTTCTGCTGGAGGTGGCAGCTGGCCATCCCCACCCGCC  900
841

907   GAGCTGAACAAGGCCACCCAGGACTTCTGCTGGAGGTGCAGCTGGGCATCCCCACCGCGTG  966
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GAGCTGAACAACGCGCACCCAGGACTTCTGCTGGAGGTGCAGCTGGGCATCCCCACCGCGTG  960
901

967   GGCCTGGAAGAAGCGCCACCCAGGCCTTCGGAGGGCGCACGCCTCACTTCAGCGTG  1026
Qy    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    GGCCTGGAAGAAGCGCCACCCAGGCCTTCGGAGGGCGCACGCCTCACTTCAGCGTG  1020
961

1027  CCCCTGGACGGAGCCTTCGGCGATGTCACCAGTACACGGCCTTCAACCCCGAGATC  1086
Qy    |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    CCCCTGGACGGAGCCTTCGGCGATGTCACCAGTACACGGCCTTCAACCCCGAGATC  1080
1021

1087  ACCCCGCCGGCATCCGCTACCAGTACAACCGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC  1146
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    ACCCCGCCGGCATCCGCTACCAGTACAACGGTGCTCAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC  1140
1081

1147  ATCTTCCAGAGCAGCAGCAAGATCCTCGAGACCCGCCGCAACCCCGAGATC  1206
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    ATCTTCCAGAGCAGCAGCAAGATCCTGGAGACCCGCCGCAACCCCGAGATC  1200
1141

1207  GTGATCTACCAGGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCAGCACCGCGCC  1266
Qy    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2347 GACGGCCATCGATGGGGCATCGTGATCTACCAGTACATGAGCGACGACCTGTACGTGGGCCAGC 2406
          ||||||||||||||||||||||||||||||    |||||||||||||||||||||||||||||||
Db   2341 GACGGCATCGATGGGGCATCGTGATCGATGATCTACCAGTACATGAGCGACGACCTGTACGTGGGCCAGC 2400

Qy   2407 GGCGGCCCTTAGGATCGATGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GGCGGCCCTTAGGATCGATGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2445
```

RESULT 2
US-10-190-305A-37

; Sequence 37, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 37
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: p2Pol.opt.YMWM_C
US-10-190-305A-37

Query Match   99.4%; Score 2443.4; DB 14; Length 2445;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2444; Conservative 0; Mismatches 0; Indels 1; Gaps 0;
```

```
Qy      7 GCCACCATGGCCGAGGCCATGAGCCAGGCCAACATCCTGATCCAGCAGGCCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGAGGCCATGAGCCAGGCCAACATCCTGATCCAGCAGGCCAGC 60

Qy     67 AACTTCAAGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 120

Qy    127 GCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC 186
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC 180

Qy    187 CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 240

Qy    247 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACCAGGGCCAACAGCCCCACCAGCCGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACCAGGGCCAACAGCCCCACCAGCCGC 300

Qy    307 GCCCATCGAGCTGCGCGGCCAGGCGGGCGGGCCGGCCCGCAGCGGCCACC 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCCCATCGAGCTGCGCGGCCAGGCGGGCGGGCCGGCCCGCAGCGGCCACC 360

Qy    367 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGACGGTGAAGGTGGGGCGGC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGACGGTGAAGGTGGGGCGGC 420

Qy    427 CAGATCAAGGAGGAGGGCCCTGCTGGACACACCGGCACCACCGTGCTGGAGGAGATGAGC 486
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGGAGGGCCCTGCTGGACACACCGGCACCACCGTGCTGGAGGAGATGAGC 480

Qy    487 CTGCCCGGCCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGGCTTCATCAAGGTGCGC 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGCCCGGCCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGGCTTCATCAAGGTGCGC 540

Qy    547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGCTGGTGGGCCAGC 606
          ||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db    541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGCTGGTGGGCCAGC 600

Qy    607 GGCCCCACCACCCCGGTGAACATCGGCCGCAACATGTCCCGTGAACATGTCCCGTGCTGGTGCACCCTG 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCCCACCACCCCGGTGAACATCGGCCGCAACATGTCCCGTGCTGGTGCACCCTG 660

Qy    667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC 720

Qy    727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGC 780

Qy    787 GAGGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACAACACC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACAACACC 840

Qy    847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 900

Qy    907 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC 960

Qy    967 GGGCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCCTACTTCAGCGTG 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGGCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCCTACTTCAGCGTG 1020

Qy   1027 CCCCTGGACGAGGACTTCCGCAAGTACAACCGCCTTCACCATCCCCAGCGATCAACAACGAG 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGAGGACTTCCGCAAGTACAACCGCCTTCACCATCCCCAGCGATCAACAACGAG 1080

Qy   1087 ACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1146
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1140

Qy   1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1206
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1200

Qy   1207 GTGATCTACCAGGGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCC 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGATCTACCAGGGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCC 1260

Qy   1267 AAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAAG 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAAG 1320

Qy   1327 CACCAGAAGAGACCCCCTTCCTGCCCATGACCTGAGCTGCACCCCGACAAGTGACCGTGCAG 1386
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CACCAGAAGAGACCCCCTTCCTGCCCATGACCTGAGCTGCACCCCGACAAGTGACCGTGCAG 1380

Qy   1387 CCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGGC 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGGC 1440

Qy   1447 AAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTG 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTG 1500

Qy   1507 CTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGATGAGC 1566
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGATGAGC 1560

Qy   1567 CTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGC 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGC 1620
```

```
;    SOFTWARE: PatentIn Ver. 2.0
;    SEQ ID NO 9
;    LENGTH: 3930
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmut_C
US-10-190-435-9

  Query Match       99.2%;   Score 2436.4;   DB 14;  Length 3930;
  Best Local Similarity  100.0%;   Pred. No. 0;
  Matches 2437;   Conservative   0;  Mismatches   0;  Indels   1;  Gaps   0;

Qy      14 TGGCCGAGCCATGAGCCAGGCCACCAGCATCCTGATGCAGCGCCAACTTCA 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1487 TCGCCGAGCCATGAGCCAGGCCACCAGCATCCTGATGCAGCGCCAACTTCA 1546

Qy      74 AGGGCCCCAAGCGGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 133
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1547 AGGGCCCCAAGCGGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 1606

Qy     134 ACTGCCGGCGCCCCCCGCAAGAGAGGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1607 ACTGCCGGCGCCCCCCGCAAGAGAGGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 1666

Qy     194 AGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCA 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1667 AGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCA 1726

Qy     254 AGGCCCGGCGAGTTCCCCAGCGAGCAGAACCCGGCCCACAGCCCACCAGCGCGAGCTGC 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1727 AGGCCCGGCGAGTTCCCCAGCGAGCAGAACCCGGCCCACAGCCCACCAGCGCGAGCTGC 1786

Qy     314 AGGTGCGCGGCGACAACCCCCGCAGCGAGCGCCCCCTGGTGCAGCGCCAGCGATCATCAAGTGGACAGAGATCATC 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1787 AGGTGCGCGGCGACAACCCCCGCAGCGAGCGCCCCCTGGTGCAGCGCCAGCGATCATCAAGTGGACAGAGATCATC 1846

Qy     374 TCCCCAGATCACCCTGTGCGACACCGGCGGCGGCCCCGACGACACCGTGCTGGAGGAGGATGAGCCTGCCCG 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1847 TCCCCAGATCACCCTGTGCGACACCGGCGGCGGCCCCGACGACACCGTGCTGGAGGAGGATGAGCCTGCCCG 1906

Qy     434 AGGAGGCCCTGCTGATCAAGAGAGGAGAAGGTTACCTGAGCTGGGTGCCCCGCCCCACAAGGGCATC 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1907 AGGAGGCCCTGCTGATCAAGAGAGGAGAAGGTTACCTGAGCTGGGTGCCCCGCCCCACAAGGGCATC 1966

Qy     494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCCGCAAGAGGTGCGCCAGTACG 553
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCCGCAAGAGGTGCGCCAGTACG 2026

Qy     554 ACCAGATCCTGACAATCTGCGGCAACATGCTGACCCAGTGGCTGCACCCTGAACTTCC 613
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2027 ACCAGATCCTGACAATCTGCGGCAACATGCTGACCCAGTGGCTGCACCCTGAACTTCC 2086

Qy     614 CCCCGTGACACATCATCGGCCGCAACATCGGCCAGCTGAAGCCCCCA 673
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2087 CCCCGTGACACATCATCGGCCGCAACATCGGCCAGCTGAAGCCCCCA 2146

Qy     674 CCATCAGCCCATGGCCCCTGACCGGCGCCATCTGCGAGGAGA 733
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2147 CCATCAGCCCATGGCCCCTGACCGGCGCCATCTGCGAGGAGA 2206

Qy     734 TGAAGCAGTGCCCCTGACCGGAGAAGATCAAGGCCCTGAGGCCCTGACCGGAGGA 793
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2207 TGAAGCAGTGCCCCTGACCGGAGAAGATCAAGGCCCTGACCGGAGGA 2266

Qy     794 TGGAGAAGGAGGGCAAGATCACCAGACACCCCGTGT 853
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2267 TGGAGAAGGAGGGCAAGATCACCAGACACCCCGTGT 2326

Qy     854 TCGCCATCAAGAAGAGGACAGCACCAAGTCGGCGAAGCTTCCGGAGCTGA 913
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2327 TCGCCATCAAGAAGAGGACAGCACCAAGTCGGCGAAGCTTCCGGAGCTGA 2386

Qy    1627 AAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGCAGTGACCTACCAGGATCTAC 1686
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 AAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGCAGTGACCTACCAGGATCTAC 1680

Qy    1687 CAGGAGGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACACC 1746
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 CAGGAGGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACACC 1740

Qy    1747 AACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATC 1806
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 AACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATC 1800

Qy    1807 TGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGACCTGGTGG 1866
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 TGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGACCTGGTGG 1860

Qy    1867 ACCGACTACTGGCAGGCCACCTGGAGTTCGTGAACACCCCCCTG 1926
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 ACCGACTACTGGCAGGCCACCTGGAGTTCGTGAACACCCCCCTG 1920

Qy    1927 GTGAAGCTGTGTGGTACCAGCTGGTGAGAGAGGAGCCCATCATCGCCGCCGAGACCTTCTACGTG 1986
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 GTGAAGCTGTGTGGTACCAGCTGGTGAGAGAGGAGCCCATCATCGCCGCCGAGACCTTCTACGTG 1980

Qy    1987 GACGGCGCCCGCAACCCGAGACCAAGCACCGGCCAAGGCCGGCTACCGGCCGACCGGGGC 2046
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 GACGGCGCCCGCAACCCGAGACCAAGCACCGGCCAAGGCCGGCTACCGGCCGACCGGGGC 2040

Qy    2047 CGGCAGAAGATCGTGACCGACGGCAGCGAGGTGAACATGTGACCGACCAGCGCCAGTGCCGAGCTGCAGGCCATC 2106
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 CGGCAGAAGATCGTGACCGACGGCAGCGAGGTGAACATGTGACCGACCAGCGCCAGTGCCGAGCTGCAGGCCATC 2100

Qy    2107 CAGCTGGCCCTGCAGGACAGCGGCCTGGAGGTGAACATCGTGACCGACCAGCAGCGCCAGTACGCC 2166
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 CAGCTGGCCCTGCAGGACAGCGGCCTGGAGGTGAACATCGTGACCGACCAGCAGCGCCAGTACGCC 2160

Qy    2167 CTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATC 2226
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 CTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATC 2220

Qy    2227 GAGCAGCTGATCAAGAGAGGAGAAGGTTACCTGAGCTGGGTGCCCGCCCCACAAGGGCATC 2286
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 GAGCAGCTGATCAAGAGAGGAGAAGGTTACCTGAGCTGGGTGCCCGCCCCACAAGGGCATC 2280

Qy    2287 GGCGGCAACGACGCAGATCGACAACTGTGTCGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTG 2346
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GGCGGCAACGACGCAGATCGACAACTGTGTCGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTG 2340

Qy    2347 GACGGCGATCGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGC 2406
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GACGGCGATCGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGC 2400

Qy    2407 GGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 GGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2445
```

RESULT 3
US-10-190-435-9
;   Sequence 9, Application US/10190435
;   Publication No. US20030143248A1
;   GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Estrelita J.
;   APPLICANT: VAN RENSBURG, Susan
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319

```
Qy   914  ACAAGCGCACCCAGACTTCTGGGAGGTCAGCTGGCATCCCCACCCGCCGGCCTGA  973
          ||||||||||||| |||||||||||||| ||||||||||||||||||||||||||
Db  2387  ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCACCCGCCGGCCTGA 2446

Qy   974  AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTCCCCTGG 1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2447  AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTCCCCTGG 2506

Qy  1034  ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACACGAGACCCCCG 1093
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2507  ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACACGAGACCCCCG 2566

Qy  1094  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC 1153
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  2567  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC 2626

Qy  1154  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCCCGCAACCCCGAGATCGTGATCT 1213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2627  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCCCGCAACCCCGAGATCGTGATCT 2686

Qy  1214  ACCAGGCCCCCCTGTACGTGGGCCAGCGACCTGGAGATCGGCCAGCACCGGCCCAAGATCG 1273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2687  ACCAGGCCCCCCTGTACGTGGGCCAGCGACCTGGAGATCGGCCAGCACCGGCCCAAGATCG 2746

Qy  1274  AGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAAGCACCAGA 1333
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2747  AGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAAGCACCAGA 2806

Qy  1334  AGGAGCCCCCTTCTGCACCCGACTGCAAGTGACCGTGCAGCCCATCG 1393
          |||||||||||||||||||||||||||||||||||||||||||||||
Db  2807  AGGAGCCCCCTTCTGCACCCGACTGCAAGTGACCGTGCAGCCCATCG 2866

Qy  1394  AGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 1453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2867  AGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 2926

Qy  1454  ACTGGGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCG 1513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2927  ACTGGGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCG 2986

Qy  1514  GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCG 1573
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2987  GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCG 3046

Qy  1574  AGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACC 1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3047  AGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACC 3106

Qy  1634  TGGTGGCCGAGATCCAGAAGCAGGGGCCACGACCAGTGGACTGACCTACCAGGAGAGC 1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3107  TGGTGGCCGAGATCCAGAAGCAGGGGCCACGACCAGTGGACTGACCTACCAGGAGAGC 3166

Qy  1694  CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCGGCACCGGCCACCAACGACG 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3167  CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCGGCACCGGCCACCAACGACG 3226

Qy  1754  TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCA 1813
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3227  TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCA 3286

Qy  1814  AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGAACGACCTGGTGGACCGACT 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3287  AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGAACGACCTGGTGGACCGACT 3346

Qy  1874  ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCTGGTGAAGC 1933
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3347  ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCTGGTGAAGC 3406

Qy  1934  TGTGGTACCAGCTGGAAGAAGAGCGCCATCATCGGCGCCGAGACTTCACGTTGGACCGGCG 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3407  TGTGGTACCAGCTGGAAGAAGAGCGCCATCATCGGCGCCGAGACTTCACGTTGGACCGGCG 3466

Qy  1994  CCGGCAACCGCGAGACCAAGATCGGCAGCCCGCTCACGTGACCGAGCCGGGCCGGCAGA 2053


Db  3467  CCGCCAACCGCGAGACCAAGATCGGCAGCCCGCTACGTGACCGACCGGGCCGGCAGA 3526
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2054  AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 2113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3527  AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 3586

Qy  2114  CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 2173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3587  CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 3646

Qy  2174  TCATCCAGGCCCAGCCCGACAGAAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGC 2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3647  TCATCCAGGCCCAGCCCGACAGAAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGC 3706

Qy  2234  TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 2293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3707  TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 3766

Qy  2294  ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 2353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3767  ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 3826

Qy  2354  TCGATGTGGGCATCGTGATCTACCAGTACATGACGAACCTGTACGTGGGCAGCGGCGCC 2413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3827  TCGATGTGGGCATCGTGATCTACCAGTACATGACGAACCTGTACGTGGGCAGCGGCGCC 3886

Qy  2414  CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
          ||||||||||||||||||||||||||||||||||||||
Db  3887  CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 3924
```

```
Query Match             99.1%;  Score 2434.8;  DB 14;  Length 3930;
Best Local Similarity   99.9%;  Pred. No. 0;
Matches 2436;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    14  TGGCCCAGGCCATGAGCGCCAGCCACCAGCCCAACATCCTGATGCAGCGCCAGAACTTCA  73
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1487  TCGCCCGAGGCCATGAGCGCCAGCCACCAGCCCAACATCCTGATGCAGCGCCAGAACTTCA 1546

Qy    74  AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1547  AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 1606

Qy   134  ACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1607  ACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 1666
```

```
Qy    194  AGGACTGCACCGAGCGCCAGCGCCAACTTCTTCCGCGAGACCTGGCCTTCCCCAGGGCA   253
           ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   1667  AGGACTGCACCGAGCGCCAGCGCCAACTTCTTCCGCGAGACCTGGCCTTCCCCAGGGCA  1726

Qy    254  AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGGAGCTGC   313
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727  AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGGAGCTGC  1786

Qy    314  AGGTGCGGCGGCACAACCCCGCAGCGAGCAGCGAGCCGGCCCGAGCGCCCAGGGCACCCTGAACT   373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787  AGGTGCGGCGGCACAACCCCGCAGCGAGCAGCGAGCCGGCCCGAGCGCCCAGGGCACCCTGAACT  1846

Qy    374  TCCCCCAGATCACCCCTGTGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCA   433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847  TCCCCCAGATCACCCCTGTGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCA  1906

Qy    434  AGGAGGCCCTGCTGGACACCGGCGGCCGGCGCCGGAGCACCGTGCTGGAGGAGATGAGCCTGCCCG   493
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907  AGGAGGCCCTGCTGGACACCGGCGGCCGGCGCCGGAGCACCGTGCTGGAGGAGATGAGCCTGCCCG  1966

Qy    494  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCGGCACCGTGCTGATCGGCCCCA   553
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCGGCACCGTGCTGATCGGCCCCA  2026

Qy    554  ACCAGATCCTGATCGAGATCTGCGGCGAGAGGCCATCGGCAGCCGCCCTGCTGACCCTGGGCTGCTGCACCCTGAACTTCC   613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027  ACCAGATCCTGATCGAGATCTGCGGCGAGAGGCCATCGGCAGCCGCCCTGCTGACCCTGGGCTGCTGCACCCTGAACTTCC  2086

Qy    614  CCCCGTGAACATCATCGGCCGCAACATGTCTGACCCAGTTGGGCTTGCACCCTGAACTTCC   673
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2087  CCCCGTGAACATCATCGGCCGCAACATGTCTGACCCAGTTGGGCTTGCACCCTGAACTTCC  2146

Qy    674  CCATCAGCGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGAGACGGACGGCCCCAAGG   733
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2147  CCATCAGCGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCTGAAGCCCGAGACGGACGGCCCCAAGG  2206

Qy    734  TGAAGCAGTGGCCCTGACCAGGAAGATCAAGGCCCCTGACCCCCTACTACAACACCCCCGTGT   793
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2207  TGAAGCAGTGGCCCTGACCAGGAAGATCAAGGCCCCTGACCCCCTACTACAACACCCCCGTGT  2266

Qy    794  TGGAGAAGGAGGCAGCAGCACGACAGCAGCACCAAGGAGATCGGCGTGCAAGCTTCGGCAAGACCCCGTGT   853
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2267  TGGAGAAGGAGGCAGCAGCACGACAGCAGCACCAAGGAGATCGGCGTGCAAGCTTCGGCAAGACCCCGTGT  2326

Qy    854  TCGCCCATCAAGAGAGGAGTACCAAGTGCGCAAGCTGGTGGCAAGGCACCAAGGTGTGGACTTCGGCGAGCTGA   913
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327  TCGCCCATCAAGAGAGGAGTACCAAGTGCGCAAGCTGGTGGCAAGGCACCAAGGTGTGGACTTCGGCGAGCTGA  2386

Qy    914  ACAAGCGCACCCAGGACTTCTGGACGTGCAGCTGGGCATCCCCACCCCCGGCCTGA   973
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387  ACAAGCGCACCCAGGACTTCTGGACGTGCAGCTGGGCATCCCCACCCCCGGCCTGA  2446

Qy    974  AGAAGCAGAAGAGCTTCCGCCAAGTACCAGCGCCTGACGCCCTACTTCAGCGTGCCCCTGG   1033
           ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   2447  AGAAGCAGAAGAGCTTCCGCCAAGTACCAGCGCCTGACGCCCTACTTCAGCGTGCCCCTGG  2506

Qy   1034  ACGAGGCCCCCCCTGTACGTGCTGACCCGCCTTCACCATCCCCAGCATCAACAAGGAGACCCCCG   1093
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2507  ACGAGGCCCCCCCTGTACGTGCTGACCCGCCTTCACCATCCCCAGCATCAACAAGGAGACCCCCG  2566

Qy   1094  GCATCCGCTACCAGTACAACGTGCTGCCGCGCCCCACCCCCAGCATCCCCAGCATCTTCC   1153
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2567  GCATCCGCTACCAGTACAACGTGCTGCCGCGCCCCACCCCCAGCATCCCCAGCATCTTCC  2626

Qy   1154  AGAGCAGCAGATGACCAAGATCCTTGGAGCCCTTCCGCGCCGCCAACCCGACATCGTGATCT   1213
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627  AGAGCAGCAGATGACCAAGATCCTTGGAGCCCTTCCGCGCCGCCAACCCGACATCGTGATCT  2686

Qy   1214  ACCAGGCCCCCCTGTACGTGGGACCTGAGATCGGCCGCCGCCAGATCGAGATCGGCCGCCAGATCG   1273
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2687  ACCAGGCCCCCCTGTACGTGGGACCTGAGATCGGCCAGATCGGCCGCCAGATCG  2746

Qy   1274  AGGAGCTGCGCAAGCTGCTGCGCGTGGGGGTTCACCACCCCCGGACAAGAAGCACCAGA   1333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2747  AGGAGCTGCGCAAGCACCTGCTGCGCGTGGGGGTTCACCACCCCCGACACAGAAGCAGAAGCACCAGA  2806

Qy   1334  AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCGACAGTGCAGCGTGGACCGTG   1393
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2807  AGGAGCCCCCCTTCCCTGCCCATCGAGACTGCGCCACCCCGACAGTGCAGCGTGGACCCCATCG  2866

Qy   1394  AGCTGCCCGAGAAGGAGGAGTCGGACCTGACCGTGAGCATCGACGAGCGATGCGGGCCAGCTGA   1453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2867  AGCTGCCCGAGAAGGAGGAGTCGGACCTGACCGTGAGCATCGACAGCGATGCGGGCCAAGCTGA  2926

Qy   1454  ACTGGGCCAGATCGAGCTACCCGTGAGCATCAAGGTGGCCCAGCTGTCGCAGCTGTGCTGCGCG   1513
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2927  ACTGGGCCAGATCGAGCTACCCGTGAGCATCAAGGTGGCCCAGCTGTCGCAGCTGTGCTGCGCG  2986

Qy   1514  GCGCCAAGGCCCTGACCGACATCGTGCCCGTGAGCCCGTGCACGGCCGTGCTGACTACGACCCCAGCAAGGACC   1573
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2987  GCGCCAAGGCCCTGACCGACATCGTGCCCGTGAGCCCGTGCACGGCCGTGCTGACTACGACCCCAGCAAGGACC  3046

Qy   1574  AGAACCCGCGAGTCCTGCGCGAGAGCAGGGCCAGGGCCCACGACCTGTACTACGACCCCAGCAAGGAGC   1633
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3047  AGAACCCGCGAGATCCTGCGCGAGAGCAGGGCCAGGGCCCACGACCTGTACTACGACCCCAGCAAGGAGC  3106

Qy   1634  TGGTGGCCGAGATCCAGAAGGCCCAAGATGCCATCGGCGGCTTCATCGGCACCGTGACCTACTACCAGGAGC   1693
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3107  TGGTGGCCGAGATCCAGAAGGCCCAAGATGCCATCGGCGGCTTCATCGGCACCGTGACCTACTACCAGGAGC  3166

Qy   1694  CCTTCAAGAACCTGACCCAGTTGGGCTTGCACCCTGAACTTCC   1753
           |||||  |||||||  |||||||||||||||||||||||||||
Db   3167  CCTTCAAGAACCTGACCCAGTTGGGCTTGCACCCACGACGACG  3226

Qy   1754  TGAAGCAGCTGACCGAGAGATGCCATGGTGCAGAAGATGCCGAGAGATCTGGGCGGCA   1813
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   3227  TGAAGCAGCTGACCGAGAGATGCCATGGTGCAGAAGATGCCGAGATCTGATCTGGGGCA  3286

Qy   1814  AGACCCCCAAGTTCCGCGATCCCCGAGTCCCGAGTTCGTGAACACACCCCCCGTGAAGCGACT   1873
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3287  AGACCCCCAAGTTCCGCGATCCCCGAGTCCCGAGTTCGTGAACACACCCCCCGTGGACCGACT  3346

Qy   1874  ACTGGCGCCACCTGACTCCCGAGTGGAGTTCGTGAACACCCCCCTGGTGAAGC   1933
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3347  ACTGGCGCCACCTGACTCCCGAGTGGAGTTCGTGAACACCCCCCTGGTGAAGC  3406

Qy   1934  TGTGGTACCAGCGGACAGCGGCAGCGGGAGCTTCGGCAAGATCGGCGGCG   1993
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3407  TGTGGTACCAGCGGACAGCGGCAGCGGGAGCTTCGGCAAGATCGGCGGCG  3466

Qy   1994  CGCCCAACCGCGAGACCAAGATCGGCGAGACCAAGATCGGCCGCGCTACGTGACCGGGCCAGA   2053
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3467  CGCCCAACCGCGAGACCAAGATCGGCGAGACCAAGATCGGCCGCTACGTGACCGGCCGGCAGA  3526

Qy   2054  AGATCGTGAGCCTGACCGGCATCGGCAGCGAGACCAAGATCGCAGGGCCATCCAGCTGG   2113
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3527  AGATCGTGAGCCTGACCGGCATCGGCAGCGAGACCAAGATCGCAGGGCCATCCAGCTGG  3586

Qy   2114  CCCTGCAGGACAGCGGCAGCGGCGGCCGGCCGGCAGCGAGTGAACATCGTGACCGGAGGTGACAGCAGC   2173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3587  CCCTGCAGGACAGCGGCAGCGGCGGCCGGCCGGCAGCGAGTGAACATCGTGACCGGTACGCGCAGCAGC  3646

Qy   2174  TCATCCAGGCCCAGCCCGACAACGTGCTGCCGCGCCCCACCCCCGACAAGGGCATCAGCGAGCAGC   2233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3647  TCATCCAGGCCCAGCCCGACAACGTGCTGCCGCGCCCCACCCCCGACAAGGGCATCAGCGAGCAGC  3706

Qy   2234  TGATCAAGAGAGAGAGGTGTACCTGACCGCCCCCAAAGGGCATCGGCGGCA   2293
           |||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   3707  TGATCAAGAGAGAGGAGAAGGTGTACCTGAGCTGGCAAGGGCATCGGCGGCA  3766

Qy   2294  ACGAGCAGATCGACAAGTCGTGGTACAGCTGGTGGCAAGGTGCTGTTCCTGACGCGGCA   2353
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3767  ACGAGCAGATCGACAAGTCGTGGTACAGCTGGTGGCAAGGTGCTGTTCCTGACGCGGCA  3826

Qy   2354  TCGATGGCCGCATCCTGATCTACCAGTACATGACGACCTTGTACTGTGGCAAGAAGCACCAGA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           ...
```

```
RESULT 5
US-10-190-435-11
; Sequence 11, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutina_C
US-10-190-435-11

Query Match    99.1%;  Score 2434.8;  DB 14;  Length 3930;
Best Local Similarity 99.9%;  Pred. No. 0;
Matches 2436;  Conservative 0;  Mismatches 2;  Indels 2;  Gaps 0;
```

```
Qy     14 TGGCCGGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCGCAACTTCA 73
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 TGGCCGGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCGCAACTTCA 1546

Qy     74 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 133
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 1606

Qy    134 ACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA 193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 ACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA 1666

Qy    194 AGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667 AGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 1726

Qy    254 AGGCCCCGGAGTTCCCCAGCGAGCAGAGACCGCGCCAACAGCCCCACCAGCCGCGGAGCTGC 313
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727 AGGCCCCGGAGTTCCCCAGCGAGCAGAGACCGCGCCAACAGCCCCACCAGCCGCGGAGCTGC 1786

Qy    314 AGTGCGGCGGCGACAACCCCCGAGCCCCGAGCGGCGCCAGGGCACCCTGAACT 373
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787 AGTGCGGCGGCGACAACCCCCGAGCCCCGAGCGGCGCCAGGGCACCCTGAACT 1846

Qy    374 TCCCCCAGATCACCCTGTGGCAGCGCGCCCCCTGTGACCATCAAGGTGGGCCGCCAGATCA 433
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847 TCCCCCAGATCACCCTGTGGCAGCGCGCCCCCTGTGACCATCAAGGTGGGCCGCCAGATCA 1906

Qy    434 AGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTCCTGGAGGAGATGAGCCTGCCCG 493
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907 AGGAGGCCCTGCTGGACACCGGCGCCGACACACCGTCCTGGAGGAGATGAGCCTGCCCG 1966

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 2026

Qy    554 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCGGCCTGCGGCCC 613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCGGCCTGCGGCCC 2086

Qy    614 CCCCGTGAACTA...
Db   2087 CCCCGTGAACATCATCG...

Qy    674 CCATCAGCCCATCGAGACCGTGCCCGTGAAGTGCCATGACGGCCCAAGG 733
Db   2147 CCATCAGCCCATCGAGACCGTGCCCGTGAAGTGCCATGACGGCCCAAGG 2206

Qy    734 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCGAGGAGA 793
Db   2207 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGAGATCTGCGAGGAGA 2266

Qy    794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCGGAACCCTACACACCCCGTGT 853
Db   2267 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCGGAACCCTACAAACACCCCGTGT 2326

Qy    854 TCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
Db   2327 TCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 2386

Qy    914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 973
Db   2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 2446

Qy    974 AGAGAAGGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCGTTACTTCAGCGTGCCCCTGG 1033
Db   2447 AGAGAAGGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTTACTTCAGCGTGCCCCTGG 2506

Qy   1034 ACGAGGACTTCCGCAAGTACACCGGCCTTCAACATCCCAGATCAACAACGAGACCCCCG 1093
Db   2507 ACGAGGACTTCCGCAAGTACACCGGCCTTCACTATCCCAGCATCAACAACGAGACCCCCG 2566

Qy   1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 1153
Db   2567 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 2626

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAATCGGCCCGCA 1213
Db   2627 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGATCGTGATCT 2686

Qy   1214 ACCAGGCCCCCCTCGTACGTGGGCAGCGACCTGGAGATCGGCCAAGATCG 1273
Db   2687 ACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAAGATCG 2746

Qy   1274 AGGAGCTGGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 1333
Db   2747 AGGAGCTGGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 2806

Qy   1334 AGGAGCCCCCTTCCTGCCCATGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCG 1393
Db   2807 AGGAGCCCCCTTCCTGCCCATGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCG 2866

Qy   1394 AGTCGCCCGAGAAGGAGCGAGCTGGACGCCACCGTGAACACATCCAGAAGCTGGTGGGCAAGCTGA 1453
Db   2867 AGTGCCCGAGAAGGAGGAGCTGGACGCCACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 2926

Qy   1454 ACTGGGCCAGCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGC 1513
Db   2927 ACTGGGCCAGCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGC 2986

Qy   1514 GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCCGAGCTGGCGGCCC 1573
Db   2987 GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCCGAGCTGGCGGCCC 3046

Qy   1574 AGAACCGCGAGATCTTGCGCGAGCCCGTGCTACTAGACCCCAGCAAGGACC 1633
Db   3047 AGAACCGCGAGATCCTGCGCGAGCCCGTGCTACTAGACCCCAGCAAGGACC 3106

Qy   1634 TGGTGGCCGAGATCCAGAAGCAGGGGCCACCAGTGGCGACCTACCAGTGCCTGATCGGCCCCA 1693
```

```
;   LENGTH: 5184
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: TatRevNefgagCpolna C
US-10-190-435-58

Query Match          99.1%;    Score 2434.8;    DB 14;    Length 5184;
Best Local Similarity 99.9%;    Pred. No. 0;
Matches 2436;  Conservative 0;  Mismatches 0;  Indels 2;  Gaps 0;

Qy     14 TCGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCGCAGCAACTTCA    73
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 TCGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCGCAGCAACTTCA  2800

Qy     74 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA   133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA  2860

Qy    134 ACTGCCGCGCCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGAGGGCCACCAGGATGA    193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 ACTGCCGCGCCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGAGGGCCACCAGAGATGA  2920

Qy    194 AGGACTGCACCACGAGGCCAACTTCTTCCGCGAGGACCTGCCTTCCCCAGGGCA    253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGACTGCACCACGAGGCCAACTTCTTCCGCGAGGACCTGCCTTCCCCAGGGCA  2980

Qy    254 AGCCCGGAGTTCCCGGCCGAGCAGAACCCGCCGCGCCGGCGCCCAACAGCCGGCGAGCTGC   313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 AGCCCGGAGTTCCCGGCCGAGCAGAACCCGCCGCGCCGGCGCCCAACAGCCGGCGAGCTGC  3040

Qy    314 AGGTGCGCGGCGACAACCCCGCAGCGAGGCGCCGGCCAGGGCACCCTGAACT    373
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGCGGCGACAACCCCGCAGCGAGGCGCCGGCCAGGGCACCCTGAACT  3100

Qy    374 TCCCCCAGATCACCCTGTGCGACATCACCCTGTGAGCATCAAGGTGGGCCGCCAGATCA    433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCCAGATCACCCTGTGCGACATCACCCTGTGAGCATCAAGGTGGGCCGCCAGATCA  3160

Qy    434 AGAGAGGGCCCTGCCTGGACACCGGCCGACGACACCGTGCTGCTGGAGCCTGAGGCCTGCCG    493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGAGAGGGCCCTGCCTGGACACCGGCCGACGACACCGTGCTGCTGGAGCCTGAGGCCTGCCG  3220

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGGCCATCGGCAAGAGGCCATCGGCCTTCATCAAGGTGGCCCAGTACG    553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATCGGCGGCCATCGGCAAGAGGCCATCGGCCTTCATCAAGGTGCGCCAGTACG  3280

Qy    554 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGGCCATCGGCACCGTGCTGTCTGATCGGCCCCA    613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGGCCATCGGCACCGTGCTGTCTGATCGGCCCCA  3340

Qy    614 CCCCCGTGAACATCATCGGCCCGCAACTGCTGACCCAGCTGGCTGACACCTGGCTCC    673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCCGTGAACATCATCGGCCCGCAACTGCTGACCCAGCTGGGCTGCACCTGAACTTCC  3400

Qy    674 CCATCAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGCATGAGGCCCCAAGG    733
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGCATGACGGCCCCAAGG  3460

Qy    734 TGAAGCAGTGCCCCTGACCGGAGGAAGAGATCAAGGGCCTGACCGGCCATCTGCGAGGAGA    793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGCCCCTGACCGGAGGAAGAGATCAAGGGCCTGACCGGCCATCTGCGAGGAGA  3520

Qy    794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCCTACCAGAAACCCCTGT    853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT  3580

Qy    854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA    913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  3640

Qy    914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACGGCCGGCCTGA    973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   3107 TGGTGGCCAGGATCCAGAGCAGGCAGGGCCCACGACCAGTGACCTACCAGAGAGC  3166
          ||||||||||||||||||||||||||||||||||||||||||||
Qy   1694 CCTTCAAGAACCTGAAGACCGCAAGTACGCCGAAGATGGCACCGCACCAACGACG  1753

Db   3167 CCTTCAAGAACCTGAAGACCGCAAGTACGCCGAAGATGCGCCACCGCACCAACGACG  3226

Qy   1754 TGAAGCAGCTGACCGAGGCCGTGCAGAGAGATCGCCATGGAGAGATCGTGATCTGGGGCA  1813
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3227 TGAAGCAGCTGACCGAGGCCGTGCAGAGAGATCGCCATGGAGAGATCGTGATCTGGGGCA  3286

Qy   1814 AGACCCCCAAGTTCCGCCTCCTCCCATCCAGAAGGAGACCTGGGAGGACCTGGTGGACCGACT  1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3287 AGACCCCCAAGTTCCGCCTCCTCCCATCCAGAAGGAGACCTGGGAGGACCTGGTGGACCGACT  3346

Qy   1874 ACTGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGTGAAGC  1933
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3347 ACTGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGTGAAGC  3406

Qy   1934 TGTGTACCAGCTGGAGAAGGAGGACCCCATCATCGGGCGCGAGACCTTCTACGTGGACGGCG  1993
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3407 TGTGTACCAGCTGGAGAAGGAGGACCCCATCATCGGGCGCGAGACCTTCTACGTGGACGGCG  3466

Qy   1994 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGCGGCAGA  2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3467 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGCGGCAGA  3526

Qy   2054 AGATCGTGAGCGCTGAGACCGAGGACAACCACCAGAGAGACCGAGCTGCAGGCCATCCAGCTGG  2113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3527 AGATCGTGAGCGCTGAGACCGAGGACAACCACCAGAGAGACCGAGCTGCAGGCCATCCAGCTGG  3586

Qy   2114 CCCTGCAGGACAGCGGCCAGCGAGGTGAACATCGTGACCGACAGCGAGTACGCCCTGGGCA  2173
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3587 CCCTGCAGGACAGCGGCCAGCGAGGTGAACATCGTGACCGACAGCGAGTACGCCCTGGGCA  3646

Qy   2174 TCATCCAGGCCCAGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCGAGCAGC  2233
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3647 TCATCCAGGCCCAGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCGAGCACAC  3706

Qy   2234 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGCGGA  2293
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3707 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCAAAGGGCATCGCGGCA  3766

Qy   2294 ACGAGGCAGATCGACAAGCTCGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA  2353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3767 ACGAGGCAGATCGACAAGCTGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA  3826

Qy   2354 TCGATGGCGGCATCGTGATCTGTGATCTGATCTACCAGTACATGACGACCTGTATCTGGGCAGCCGGCGGCC  2413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3827 TCGATGGCGGCATCGTGATCTGTGATCTGATCTACCAGTACATGACGACCTGTATCTGGGCAGCCGGCGGCC  3886

Qy   2414 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2451
          ||||||||||||||||||||||||||||||||||||||
Db   3887 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  3924
```

RESULT 6
US-10-190-435-58
;   Sequence 58, Application US/10190435
;   Publication No. US2003014324BA1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Estrelita J.
;   APPLICANT: VAN RENSBURG, Susan
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18113.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 58

RESULT 7
US-10-190-305A-82
; Sequence 82, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 82
; LENGTH: 5184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: TatRevNefgagCpolIna C
US-10-190-305A-82

Query Match   99.1%;   Score 2434.8;   DB 14;   Length 5184;
Best Local Similarity   99.9%;   Pred. No. 0;
Matches 2436;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

```
Qy    254 AGGCCCGGAGTTCCCAGCGAGGAGCAGACAGCCCGGCAACAGCCCCACCAGCTGC 313
          ||||||||||||||||||||||||| |||||||   |||| ||| |||||||||||
Db   2981 AGGCCCGGAGTTCCCAGCGAGCAGACAGAAGCCCGAGCGCCCCACCAGCCCGGAGCTGC 3040

Qy    314 AGGTGCGGCGGCACAACCCCCGCAGCGGACCGGCGGAGCGCGCGAGGCCGGCACCCTGAACT 373
          ||||||||||||||||||||| ||||||||||||||| |||||||||||||||||||||||||
Db   3041 AGGTGCGGCGGCACAACCCCCTGCGGACCGGCGGAGCATCAAGGTGCGGCGGCACCCTGAACT 3100

Qy    374 TCCCCAGATCACCCGTGTGCAGCATCAAGGTGGCGGCCAGATCA 433
          |||||||||||||||||||| ||||||||||||||||||||||||
Db   3101 TCCCCAGATCACCCGTGTGCGAGCATCAAGGTGGCGGCCAGATCA 3160

Qy    434 AGGAGGCCCTGCTTGGACACCGGCGCCGACGACCGGCGACACCGTGCTGAGGAGAATGAGCCTGCCCG 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGCCCTGCTTGGACACCGGCGCCGACGACACCGTGCTGAGGAGAATGAGCCTGCCCG 3220

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGCTTCATCAAGGTGCGCCAGTACG 553
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATCGGCGCATCGGCTTCATCAAGGTGCGCCAGTACG 3280

Qy    554 ACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTGCTGATCGGCACCCCA 613
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTGCTGATCGGCACCCCA 3340

Qy    614 CCCCGTGAACATCGGCCGCAACATCATCGGCCTGACCTGGGCTGCACCCTGAACTTCC 673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCGTGAACATCATCGGCCTGACCTGGGCTGCACCCTGAACTTCC 3400

Qy    674 CCATCAGCCCCATCGAGAACCCTGAAGCTGAAGCCCGGCGACGCGGCCCCAAGG 733
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAGAACCCTGAAGCTGAAGCTGAAGCCCGGCGACGGGACCGCCCCAAGG 3460

Qy    734 TGAAGCAGTGCCCCTGACCGAGGAAGAATCAAGGCCCTGACCGGCCATCTGGGAGGAGA 793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGCCCCTGACCGAGGAAGAATCAAGGCCCTGACCGGCCATCTGCGGAGGAGA 3520

Qy    794 TGGGAGGAGAGGGCAAGATCACCACCGACAGCAGCAGCCCCTACACACACCCCGTGT 853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGGAGGAGAGGGCAAGATCACCACCGACAGCAGCAGCCCCTACACACACCCCGTGT 3580

Qy    854 TCGCCCATCAAGAGGTGCAGCTGCAAGTGGCGCAAGGTGGCCAGCTGGGCATCCCGCCTGA 913
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCCATCAAGAGGTGCAGCTGCAAGTGGCGCAAGGTGGCCAGCTGGGCATCCCGCCTGA 3640

Qy    914 ACAAGCGCACCCAGAGACTTCTGGGAGGTGCCAGCTGGGCATCGGGCCCTGA 973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGCGCACCCAGAGACTTCTGGGAGGTGCCAGCTGGGCATCCCCGCCCTGA 3700

Qy    974 AGAAGAAGAAGAGAGGCGTGACCGTGCTGGACCGTGGACGGGGGGCGACCGGCGCCCTGG 1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAAGAAGAAGAGAGGCGTGACCGTGCTGGACCGTGGACGGGGGGCGACCTACTTCAGCGGCCCCTGG 3760

Qy   1034 ACGACCCCCGTGTACGTGGGCAGCGACCTGGAGCGCTTCACCATCCCAGCATCAACAACGAGACCCCCG 1093
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACGACCCCCGTGTACGTGGGCAGCGACCTGGAGCGCTTCACCATCCCAGCATCAACAACGAGACCCCG 3820

Qy   1094 GCATCCGCTACCAGTACAACGGCCTGCGCGTGGGGCCTGCGCGGCAGCGCCCAGCGAGGGCAGCCAGCCAGCATCTTCC 1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGCTACCAGTACAACGGCCTGCGCGTGGGGGCTGCGCGGCAGCGCCCAGCGAGGGCAGCGCCCAGCATCTTCC 3880

Qy   1154 AGAAGCAGCAGCAAGATCCTGGAGCGCTTCCGCGGCATCGTGATCT 1213
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3881 AGAAGCAGCAGCAAGATCCTGGAGCGCTTCCGCGGCATCGTGATCT 3940

Qy   1214 ACCAGGCCCCCGTGTACGTGGGCAGCGACCTGGAGATCGGCGCCAGCACCGACTCG 1273
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3941 ACCAGGCCCCCCGTGTACGTGGGCAGCGACCTGGAGATCGGCGCCAGCACCGACATCG 4000

Qy   1274 AGGAGCTGCGCAAGGCACCTGCTGCCTGCTGGGCTTCACCACCACCCCCGACAAGAGCACCAGA 1333
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4001 AGGAGCTGCGCAAGGCACCTGCTGCCTGCTGGGCTTCACCACCACCCCCGACAAGAGCACCAGA 4060

Qy   1334 AGGAGCCCCCTTCTTGCCCATCGAGCTGCACCCGACCGCAGAGTGGACCGTCGCAGCCCATCG 1393
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   4061 AGGAGCCCCCTTCTTGCCCATCGAGCTGCACCCGACCGCAGCAAGTGGACCGTCGCAGCCCATCG 4120

Qy   1394 AGCTGCCCGAGAAGGAGAGTGGACCTGGACCGTGGACCGTGAACGACATCGAAGCGAATCAGAAGCTGGTTGGGCAAGCTGA 1453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4121 AGCTGCCCGAGAAGGAGAGTGGACCTGGACCGTGGACCGTGAACGACACATCCAGAAGCTGGTTGGGCAAGCTGA 4180

Qy   1454 ACTGGGCCAGTCGAGATCTACCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGTGTGCCGCG 1513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4181 ACTGGGCCAGTCGAGATCTACCCGGCATCAAGGTGCGCCGGCCAGCTGTGTCAAGCTGTGCTGCGCG 4240

Qy   1514 GCGGCCAAGGCCCTGACCGACAGCTGGCCCCTGACCGAGGAGGAGCCGAGCTGGAGCTGCCCG 1573
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4241 GCGGCCAAGGCCCTGACCGACACATCGTGCCCGTGCCCCTGACCGAGGAGGAGCCGAGCTGGCGCCG 4300

Qy   1574 AGAACCGGCGAGATCCTGCGCGAGATCCGCGGAGGGCGTGCTACTACGACCCCAGCAAGGACC 1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4301 AGAACCGGCGAGATCCTGCGCGAGGCCGTGCTACGACCCCAGCAAGGACC 4360

Qy   1634 TGGTGGCCGAGATCCAGAAGCAGAGGGCCACGACCAGTGGACCTACCAGGAGAGC 1693
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4361 TGGTGGCCGAGATCCAGAAGCAGAGGGCCACGACCAGTGGACCTACCAGGAGAGC 4420

Qy   1694 CCTTCAGGAACCTGAAGACCGGCCAAGTACGGCACGGCACCCACCAACGACG 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4421 CCTTCAGGAACCTGAAGACCGAAGACCGGCCAAGTACGGCACGCCCACCAACGACG 4480

Qy   1754 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGATCGTGGGCA 1813
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4481 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGATCGTGGGCA 4540

Qy   1814 AGACCCCCAAGTTCCGCACCTGATCCCGAGTTGGGAGGACCTGGGAGAGACGACT 1873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4541 AGACCCCCAAGTTCCGCACCTGATCCCGAGTTGGGAGGACCTGGGAGAGACGACT 4600

Qy   1874 ACTGGCAGGCCCACCTGATCCGGCGCGAGTTCGTGAACACCCCCCCTGGTGAAGC 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4601 ACTGGCGCGCCCACCTGATCCGGCGCGAGTTCGTGAACACCCCCCCTGGTGAAGC 4660

Qy   1934 TGTGGTACCGGGGAGGAGACCAAGATCGGCACGCCGGCCGTACGTGGACCGCG 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4661 TGTGGTACCAGCTGGAGGAGACCAAGATCGGCACGCCGGCCGTACGTGGACCGCG 4720

Qy   1994 CCGCCAACCCGAGACCAAGATCGACCAAGAGCTGACCGAGGCCGGCCGAGA 2053
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4721 CCGCCAACCCGAGACCAAGATCGACCAAGAGCTGACCGAGGCCGGCCGAGA 4780

Qy   2054 AGATCGTGACCGGCGAGCTGACCGACACCGAGAAGAGCTGCAGGCCATCCAGCTGG 2113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4781 AGATCGTGACCGGCGAGCTGACCGACACCGAGAAGAGCTGCAGGCCATCCAGCTGG 4840

Qy   2114 CCCTGCAGGACAGCGGCAGCGAGTGACATCGTGACCGACCCCTGGGCA 2173
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4841 CCCTGCAGGACAGCGGCAGCGAGTGACATCGTGACCGACCCCTGGGCA 4900

Qy   2174 TCATCCAGGCCCCAGCCCGACAAGCCCGACAAGATCATCGGGCAGC 2233
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4901 TCATCCAGGCCCCAGCCCGACAAGCCCGACAAGATCATCGAGCAGC 4960

Qy   2234 TGATCAAGAGGAGAAGGTGTACCTGACTGGAGCTGGGTGCCCCCACAAGGGCATCGGCGGCA 2293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4961 TGATCAAGAGGAGAAGGTGTACGTGAGCTGGGTGCCCCCACAAGGGCATCGGCGGCA 5020

Qy   2294 ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGCGGCA 2353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5021 ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGCGGCA 5080

Qy   2354 TCGATGGCGGCATCGTGATCGTGATCTACCAGTACGTACGGGGGCAGCGGCGCC 2413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5081 TCGATGGCGGCATCGTGATCGTGATCTACCAGTACGTACGGGGGCAGCGGCGGCC 5140

Qy   2414 CTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT 2451
```

```
Db   5141 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 5178
          |||||||||||||||||||||||||||||||||||||

RESULT 8
US-10-190-435-44
; Sequence 44, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 44
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt.YM_C

US-10-190-435-44

Query Match      98.8%;  Score 2428.6;  DB 14;  Length 2457;
Best Local Similarity  99.6%;  Pred. No. 0;
Matches 2447;  Conservative  0;  Mismatches  4;  Indels  6;  Gaps  1;
```

US-10-190-305A-38

```
Query Match             98.8%;   Score 2428.6;   DB 14;   Length 2457;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 2447;   Conservative 0;   Mismatches 4;   Indels 4;   Gaps 1;
```

```
Qy     1 GTCGACGGCACCATGGCCGAGGCCATGAGCCAGCAGGCCACCAGCGCCAACATCCTGATGCAG   60
         |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     1 GTCGACGGCCACCATGGCCGAGGCCACCAGCGCCCAGCAGGCCACCAGCGCCAACATCCTGATGCAG   60

Qy    61 CGGCAGCAACTTCAAGGGCCCAGCGCCATCATCAGCGCCCCCGGACTGGCGGCAAGGAGGGGC  120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    61 CGCGCAACTTCAAGGGCCCAGCGCCATCATCAGCGCCCCCGGACTGGCGGCAAGGAGGGC  120

Qy   121 CACATCGCCCGCAACTGCCGCGCCCCCGCAAGAGAGGCTGCTGAAGTGCGGCAAGGAG  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACATCGCCCGCAACTGCCGCGCCCCCGCAAGAGAGGCTGCTGAAGTGCGGCAAGGAG  180

Qy   181 GGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGAACAGCGGCCAACAGCCCCACC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGAACAGCGGCCAACAGCCCCACC  240

Qy   241 TTCCCCAGGGCAAGGCCGGCGAGTTCCCCAGCGGCGCGACAACCCCCGAGCGAGGCGGCCGCCAGCGGCCAG  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCCCAGGGCAAGGCCGGCGAGTTCCCCAGCGGCGCGACAACCCCCGAGCGAGGCGGCCGCCAGCGGCCAG  300

Qy   301 AGCCGCGAGCTGCAGGTGCGCGGACAACCCCCGTGCAGCGCCCCCTGGTGAGCATCAAGGTG  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCCGCGAGCTGCAGGTGCGCGGACAACCCCCGTGCAGCGCCCCCTGGTGAGCATCAAGGTG  360

Qy   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCGAGTGGTG  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCGAGTGGTG  420

Qy   421 GGCGGCCAGATCAAGGAGGCCCTGCTGCACCGGCGCCGACGACACCGTGCTGCTGGAGGAG  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGCCAGATCAAGGAGGCCCTGCTGCACCGGCGCCGACGACACCGTGCTGCTGGAGGAG  480

Qy   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGCCTTCATCAAG  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGCCTTCATCAAG  540

Qy   541 GTGCGGCCAGTACCACCAGATCTGATCGAGATCTGAGATCTGGGCAAGAAGGCCATCGGCACCCGTG  600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGCGGCCAGTACCACCAGATCTGATCGAGATCTGAGATCTGGGCAAGAAGGCCATCGGCACCCGTG  600

Qy   601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATCGGCGGCCTGAAGCTGACCAGCTGGGCTGC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATCGGCGGCCTGAAGCTGACCAGCTGGGCTGC  660

Qy   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG  720

Qy   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCC  780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCC  780

Qy   781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCAACCAAGATCGGCCCCGAGAACCCCTAC  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCAACCAAGATCGGCCCCGAGAACCCCTAC  840

Qy   841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC  900

Qy   901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC  960

Qy   961 CCCGCCGGCCTGAAGAGGAGAAGAGAGCGTGACCGTGCTGGGACGTGGGCGACCGCCTACTTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCCGGCCTGAAGAGGAGAAGAGAGCGTGACCGTGCTGGGACGTGGGCGACCGCCTACTTC 1020
```

```
Db  1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGAACCGGCAAGTACGCCAAGATGCGC 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1735 ACCGCCCACCACCAACGCGTGAAGCAGTGAAGCCGAGGCCGTCGACGAGATCGGCCATGGAG 1794
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ACCGCCCACCAACAAGACGTGAAGCAGCTGACCGAGGCCGTGACCGAGGAAGATCGGCCATGGAG 1800

Qy  1795 AGCATCGTGATCTGGGGCCAAGAGACCCCCAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1854
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGCATCGTGATCTGGGGCCAAGACCCCCAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860

Qy  1855 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1914
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920

Qy  1915 ACCCCCCCGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGCCGGT 1974
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ACCCCCCCGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGCCGAG 1980

Qy  1975 ACTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2034
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ACTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040

Qy  2035 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2094
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2100

Qy  2095 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCTGGAGGTGAACATCGTGACCGAC 2154
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCTGGAGGTGAACATCGTGACCGAC 2160

Qy  2155 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTG 2214
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTG 2220

Qy  2215 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2274
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280

Qy  2275 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2334
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2340

Qy  2335 GTGCTGTTCCTGGACGGCATCGATGGCGGCCATCGTGATCTGATCTACCAGTACATGACGACCTG 2394
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCCATCGTGATCTGATCTACCAGTACATGACGACCTG 2400

Qy  2395 TACGTGGGCAGCGGCCGGCGGCCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCCGGT 2451
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 TACGTGGGCAGCGGCCGGCGGCCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCCGGT 2457
```

RESULT 9

US-10-190-305A-38

; Sequence 38, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan
;   APPLICANT: LIAN, Ying
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
;   TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: 2302-18702 / 18702.002
;   CURRENT APPLICATION NUMBER: US/10/190,305A
;   CURRENT FILING DATE: 2002-07-05
;   NUMBER OF SEQ ID NOS: 93
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
;   LENGTH: 2457
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p2Polopt.YM_C

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCCAGCCGCCAACATCCTGATGCAGCGCAACTTCA 73
               |||||||||||||||||||||||||||||||||||||| | |||||||||||||||||||||||
Db   1088 TGGCCGAGGCGATGAGCCAGGCCACCAGCGCCCAGCCGCCAACACCAGCGTGATGATGCAGCAACTTTAAAA 1147

Qy   74 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 133
               |||||||||||||||||||||||||||||||||||||| | |||||||||||||||||||||
Db   1148 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCCGCA 1207

Qy   134 ACTGCCGGCCGCCCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
               ||||||||||||| || |||||||||||||||||||||||||||||||||||||||||||||
Db   1208 ACTGCCGGCCGCGCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 1267

Qy   194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGGAGGACCGGTGCA 253
               ||||||| || ||||||||||||||||||||||||||||||||||||
Db   1268 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGGAGGACCGGGGCA 1327

Qy   254 AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGGCCAACAGCCCCAGCGCCGCGAGCTGC 313
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   1328 AGGCCCGCGAGTTCCCCGGCGAGCAGAACCGGCCAACAGCCCCAGCGCCGCGAGCTGC 1387

Qy   1021 AGCCGTGCCCCTGGACGAGGACTTCCGCCAAGTACACCGCCTTCACCATCCCCAGCATCAAC 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGCCGTGCCCCTGGACGAGGACTTCCGCCAAGTACACCGCCTTCACCATCCCCAGCATCAAC 1080

Qy   1081 AACGAGACACCCCGGCATCCCGGCTACCAGTACAACGTCGTCTGCCCCCAGGGCTGGAAGGGCAGC 1140
               ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   1081 AACGAGACACCCCGGCATCCCGGCTACCAGTACAACGTCGTCTGCCCCCAGGGCTGGAAGGGCAGC 1140

Qy   1141 CCCAGCATCTTCCAGAGCAGCATGACGAAGATCCTGGAGCCCTTCCGCGCCCAACCCC 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCCAGAGCAGCATGACGAAGATCCTGGAGCCCTTCCGCGCCCAACCCC 1200

Qy   1201 GAGATCGTGATCTACCAGTACCGTGGGCAGCAGCCTGGGCAGCGACCTCGGAGATCGGCCAGCAC 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCAGTACCGTGGGCAGCAGCCTGGGCAGCGACCTCGGAGATCGGCCAGCAC 1260

Qy   1261 CGGCCAAGATCGAGGAGTCGCGCAAGCACCTGCGCTGCGGCTGGGGGCTTCACCACCCCCGAC 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGGCCAAGATCGAGGAGTCGCGCAAGCACCTGCGCTGCGGCTGGGGGCTTCACCACCCCCGAC 1320

Qy   1321 AAGAAGCACCAGAAGAAGGAGCCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACAAG 1374
               |||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   1321 AAGAAGCACCAGAAGAAGGAGCCCCCCTTCCTGTGTGCATGGCTACGAGCTGCACCCCGACAAG 1380

Qy   1375 TGGACCGTGCAGCCCATCGAGCTGCCCATCGAGAAGGAGAGCTGACCGTGACGACATCCAG 1434
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGCTGCCCATCGAGAAGGAGAGCTGACCGTGACGACATCCAG 1440

Qy   1435 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCCAG 1494
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCCAG 1500

Qy   1495 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1554
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy   1555 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1614
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1620

Qy   1615 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGACC 1674
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGACC 1680

Qy   1675 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGGCAAGATGCGC 1734
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGGCAAGATGCGC 1740

Qy   1735 ACCGCCCACCACCAACGACGTGAAGCAGCTGACCGAGAGATCGGCATGGAG 1794
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACCGCCCACCACCAACGACGTGAAGCAGCTGACCGAGAGATCGGCATGGAG 1800

Qy   1795 AGCATCGTGATCTGGGCCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1854
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCGTGATCTGGGCCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860

Qy   1855 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1914
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920

Qy   1915 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCCAG 1974
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCCAG 1980

Qy   1975 ACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGGCCGGCTACGTG 2034
               ||| ||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGGCCGGCTACGTG 2040

Qy   2035 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2094
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2100

Qy   2095 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC 2154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC 2160

Qy   2155 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACCAGAGCGAGAGCGAGCTGGTG 2214
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACCAGAGCGAGAGCGAGCTGGTG 2220

Qy   2215 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACCTGAGCTGGGTGCCCGCC 2274
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACCTGAGCTGGGTGCCCGCC 2280

Qy   2275 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2334
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2340

Qy   2335 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGATGGCGGCATCGATGGCGGCATGATGA 2394
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGATGGCGGCATCGATGGCGGCATGATGA 2400

Qy   2395 TACGTGGGCAGCGGCCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGGCAGCGGCCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457

```
Qy  1394  AGCTGCCCGAGAAGAGAGAGTGGACCGTGAACGACATCAGAGAAGCTGGTGGGCAAGCTGA  1453
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  2468  AGCTGCCCGAGAAGAGAGAGTGGACGTGAACGACATCAGAGAAGCTGGTGGGCAAGCTGA  2527

Qy  1454  ACTGGGCCAGCCAGATCTACCCCGGCATGCGCGCCCCTGTGAGCATCAAGGTGCGCGCCG  1513
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  2528  ACTGGGCCAGCCAGATCTACCCCGGCATGCGCGCCCCTGTGAGCATCAAGGTGCGCGCCG  2587

Qy  1514  GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCG  1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2588  GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCG  2647

Qy  1574  AGAACCCGAGATCCTGCGCGAGATCCTGCGCGAGATCCAGCCCTGCACGACCCCAGCAGGACC  1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2648  AGAACCCGAGATCCTGCGCGAGATCCTGCGCGAGATCCAGCCCTGCACTACGACCCCAGCAGGACC  2707

Qy  1634  TGGTGGCCGAGATCCAGAAGCAGAGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGC  1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2708  TGGTGGCCGAGATCCAGAAGCAGAGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGC  2767

Qy  1694  CCTTCAAGAACCTGAGACACCGGCCGTGCGAGAGATCGCTGCGTGATCTGGGGCA  1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2768  CCTTCAAGAACCTGAGACACCGGCCGTGCGAGAGATCGCTGCGTGATCTGGGGCA  2827

Qy  1754  TGAAGCAGTCGACCGAGCCGTGCGAGAGATCGCCATGCGATCTGTGACCGGACCGACT  1873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2828  TGAAGCAGTCGACCGAGCCGTGCGAGAGATCGCCATGCGATCTGTGACCGGACCGACT  2887

Qy  1814  AGACCCCCAAGTTCCGGCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACT  1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2888  AGACCCCCAAGTTCCGGCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACT  2947

Qy  1874  ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCGGTGGTGAAGC  1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2948  ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCGGTGGTGAAGC  3007

Qy  1934  TGTGGTACCAGCTGGAGAAGAGAGCCCATCATCGGCGGCCGCTACGTGGACGCGGCG  1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3008  TGTGGTACCAGCTGGAGAAGAGAGCCCATCATCGGCGGCCGCTACGTGGACGCGGCG  3067

Qy  1994  CCGCCAACCGGCGAGACCAAGATCGGCAAGGCCCGGCTACGTGACCGGCCGGCACAGA  2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3068  CCGCCAACCGGCGAGACCAAGATCGGCAAGGCCCGGCTACGTGACCGGCCGGCACAGA  3127

Qy  2054  AGATCGTGAGCCTGACCGAGCTGAGCTGAGCCAGTGCGAGGCCATCCAGCTGG  2113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3128  AGATCGTGAGCCTGACCGAGCTGAGCTGAGCCAGTGCGAGGCCATCCAGCTGG  3187

Qy  2114  CCCTGCAGGACAGCGGCGCAGCGGCAGCGAGGTGAACATCGTGACCCCAGTACGGCGGCA  2173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3188  CCCTGCAGGACAGCGGCGCAGCGGCAGCGAGGTGAACATCGTGACCCCAGTACGGCGGCA  3247

Qy  2174  TCATCCAGGCCGGCATCGTGATCTACCAGTACATGGACACCTGGTGAACCAGATCATCGAGCAGC  2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3248  TCATCCAGGCCGGCATCGTGATCTACCAGTACATGGACACCTGGTGAACCAGATCATCGAGCAGC  3307

Qy  2234  TGATCAAGAAGGAGAAGTGTACCTGAGCTGGTGCCCGCCATCGGCGGCA  2293
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3308  TGATCAAGAAGGAGAAGTGTACCTGAGCTGGTGCCCGCCATCGGCGGCA  3367

Qy  2294  ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCCGAAGTGGCTGTTCCTGGACGGCA  2353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3368  ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCCGAAGTGGCTGTTCCTGGACGGCA  3427

Qy  2354  TCGATGGCGGCATCGTGATCTACCAGTACATGGACACCTGTACGTGGGCA  2413
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3428  TCGATGGCGGCATCGTGATCTACCAGTACATGGACACCTGTACGTGGGCA  3487

Qy  2414  CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2451
          ||||||||||||||||||||||||||||||||||||||
Db  3488  CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  3525
```

```
Qy   314  AGGTGCGCGCGACAACCCCGCAGCGAGGCCGGCGCCCAGGGCACCCTGAACT   373
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388  AGGTGCGCGCGACAACCCCGCAGCGAGGCCGGCGCCCAGGGCACCCTGAACT  1447

Qy   374  TCCCCCAGATCACCCTGTGGCACGCGCCCCTGTGAGCATCAAGGTGCGGCAGATCA   433
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1448  TCCCCCAGATCACCCTGTGGCACGCGCCCCTGTGAGCATCAAGGTGCGGCAGATCA  1507

Qy   434  AGGAGGCCCTGCTGACACCGGCCGCCGGACGACCGTGCTGGAGGAGATGAGCCTGCCCG   493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1508  AGGAGGCCCTGCTGACACCGGCCGCCGGACGACCGTGCTGGAGGAGATGAGCCTGCCCG  1567

Qy   494  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCGAAGTGCGCCAGTACG   553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1568  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCGAAGTGCGCCAGTACG  1627

Qy   554  ACCAGATCCTGATCGAGATCTGCGGCCAACAGGCCATCGGCACCGTGCTGATCGGCCCCA   613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1628  ACCAGATCCTGATCGAGATCTGCGGCCAACAGGCCATCGGCACCGTGCTGATCGGCCCCA  1687

Qy   614  CCCCGTGAACATCATCGGCGCGCAACATGTCGACCAGTGGGCGTGCACCCTGAACTTCC   673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1688  CCCCGTGAACATCATCGGCGCGCAACATGTCGACCAGTGGGCGTGCACCCTGAACTTCC  1747

Qy   674  CCATCAGCCCCATCGAACCGTGCCCGTGAAGTGAAGCCCGGCGAATGGACGGCCCCAAGG   733
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1748  CCATCAGCCCCATCGAACCGTGCCCGTGAAGTGAAGCCCGGCGAATGGACGGCCCCAAGG  1807

Qy   734  TGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCCATCTGCGGAGGAGA   793
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1808  TGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCCATCTGCGGAGGAGA  1867

Qy   794  TGGGAGAAGAGAGGACGAAGATCACCAGGACTTCTGGAGGTGGCACCAAGTTCCGGCACA   853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1868  TGGGAGAAGAGAGGACGAAGATCACCAGGACTTCTGGAGGTGGCACCAAGTTCCGGCACA  1927

Qy   854  TCGCCATCAAGAAGGAGACGCACCAGGACTTCTGGGACTTCCGCGAGCTGA   913
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1928  TCGCCATCAAGAAGGAGACGCACCAGGACTTCTGGGACTTCCGCGAGCTGA  1987

Qy   914  ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCCGCCTGA   973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1988  ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCCGCCTGA  2047

Qy   974  AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGCCGACGCGCTGGACGTGCCCCTGG  1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2048  AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGCCGACGCGCTGGACGTGCCCCTGG  2107

Qy  1034  ACGAGGACTTCGGCAAGTACCACCGCCTTCACCGCCCAGCATCAACAACGAGACCCCCG  1093
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2108  ACGAGGACTTCGGCAAGTACCACCGCCTTCACCGCCCAGCATCAACAACGAGACCCCCG  2167

Qy  1094  GCATCCGCTACCAGTACAACGTGCTGCCCAGGGCCTGGAAGGGCAGCCCCAGCATCTTCC  1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2168  GCATCCGCTACCAGTACAACGTGCTGCCCAGGGCCTGGAAGGGCAGCCCCAGCATCTTCC  2227

Qy  1154  AGAAGCAGCAGATCCTGAACCAAGATCCTGGCCCTTCGCGCCCCAACCCCGAGATCGTGATCT  1213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2228  AGAAGCAGCAGATCCTGAACCAAGATCCTGGCCCTTCGCGCCCCAACCCCGAGATCGTGATCT  2287

Qy  1214  ACCAGCCCCCTGTACGTGGGCAGCGAGCTGCACCCGCCCAGCAGTGGCCCGGCCAAGATCG  1273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2288  ACCAGCCCCCTGTACGTGGGCAGCGAGCTGCACCCGCCCAGCAGTGGCCCGGCCAAGATCG  2347

Qy  1274  AGGAGCTGCCGCAAGCACCTGCTGCTGGGCGCCACCACCCCCCGACAAGAAGACCACCAGA  1333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2348  AGGAGCTGCCGCAAGCACCTGCTGCTGGGCGCCACCACCCCCCGACAAGAAGACCACCAGA  2407

Qy  1334  AGGAGCCCCCTTCCTGCCCATCGAGCTGCACCCGGACCAGTGGACCGTGCAGCCCATCG  1393
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2408  AGGAGCCCCCTTCCTGCCCATCGAGCTGCACCCGGACCAGTGGACCGTGCAGCCCATCG  2467
```

RESULT 11
US-10-190-435-45
; Sequence 45, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 45
;   LENGTH: 2457
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p2Polopt_C
US-10-190-435-45

Query Match   97.8%;   Score 2403.4;   DB 14;   Length 2457;
Best Local Similarity   99.3%;   Pred. No. 0;
Matches 2439;   Mismatches 0;   Conservative 0;   Indels 6;   Gaps 2;

Matches 2439; Conservative 0; Mismatches 6; Indels 12; Gaps 2;

```
Qy      7 GCCACCATGGCCGAGGCCATGAGCCAGCTGACCGAGAAGATCGGCCAGCGGCCCAACATCCTGATGCAGGCGAGC     66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGAGGCCATGAGCCAGCTGACCGAGAAGATCGGCCAGCGGCCCAACATCCTGATGCAGGCGAGC     60

Qy     67 AACTTCAAGGGCCCCAAGCGCATCATCAAGCTGCTTCAACTGCGGCAAGGAGGGCCAAGGAGGGCCACATC     126
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCCAAGCGCATCATCAAGCTGCTTCAACTGCGGCAAGGAGGGCCAAGGAGGGCCACATC     120

Qy    127 GCCCGCAACTGCCGCGGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGTGCGGCAAGGAGGGCCAAGGAGGGCCAC     186
            ||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCGCAACTGCCGCGGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGTGCGGCAAGGAGGGCCAAGGAGGGCCAC     180

Qy    187 CAGATGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCGCGCAAGGCCGGCTACGTG     246
            ||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCGCGCAAGGCCGGCTACGTG     240

Qy    247 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCGAGCAACCGCGCCAACAGCCCAACAGCCCACCAGCCGCGC     306
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCGAGCAACCGCGCCAACAGCCCAACAGCCCACCAGCCGCGC     300

Qy    307 GAGCTGCAGGTGCGCGGCACACCCCGCAGCGGGCGCGCCCCCTGGTGAGCGATCAAGGTGGGCGGC     366
            ||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGCTGCAGGTGCGCGGCACACCCCGCAGCGGGCGCGCCCCCTGGTGAGCGATCAAGGTGGGCGGC     360

Qy    367 CTGAACTTCCCCCAGCATCACCCGTGTGGCAGCGGCCGGACCACCGGCATCGGCGGCTTCATCAAGGTGCGC     426
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCCAGCATCACCCGTGTGGCAGCGGCCGGACCACCGGCATCGGCGGCTTCATCAAGGTGCGC     420

Qy    427 CAGATCAAGGAGGCCCTGCTGGAAGCCAAGATGATCGGCGGCCAAGAAGAGGCATCGGCGGCAAGAGGATGAGC     486
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGGCCCTGCTGGAAGCCAAGATGATCGGCGGCCAAGAAGAGGCATCGGCGGCAAGAGGATGAGC     480

Qy    487 CTGCCCGGCAAGTGAAGGCCCTGCTGGAAGCCAAGATGATCTGCGGCAAGAGGCATGCGGCAAGAAGGCCATCGGCGACCGTGCTCAAGGTGCGC     546
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGCCCGGCAAGTGAAGGCCCTGCTGGAAGCCAAGATGATCTGCGGCAAGAGGCATGCGGCAAGAAGGCCATCGGCGACCGTGCTCAAGGTGCGC     540

Qy    547 CAGTACGACCAGATCCTGATCGAGAGGCATCTGCGGCAAGAGGCCATCGGCGACCGTGGGCTGCTGCACCCTG     606
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGTACGACCAGATCCTGATCGAGAGGCATCTGCGGCAAGAGGCCATCGGCGACCGTGGGCTGCTGCACCCTG     600

Qy    607 GGCCCCACCCCGTGAACATCATCGAGACCGTGCCCGTGAAGCTGAAGCCGGCCATGAGCCCGGCATGGACCGGC     666
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCCCACCCCGTGAACATCATCGAGACCGTGCCCGTGAAGCTGAAGCCGGCCATGAGCCCGGCATGGACCGGC     660

Qy    667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGC     726
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGC     720

Qy    727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACACGGCCCATCTGC     786
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACACGGCCCATCTGC     780

Qy    787 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC     846
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC     840

Qy    847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC     906
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC     900

Qy    907 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC     966
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC     960

Qy    967 GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG     1026
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG     1020

Qy   1027 CCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAG     1086
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAG     1080
```

```
Db   1741 ACCGGCCCACCAACGACGTCGAAGCAGCTGACCGAGGCCCTGCTGCCCATCGGAGAAGATCGGCCATGGAG     1800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1795 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG     1854
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG     1860

Qy   1855 GAGACCCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC     1914
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGACCCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC     1920

Qy   1915 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGGCCGCC     1974
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGGCCGCC     1980

Qy   1975 ACTTCTACGTGGACGGCGCGCCCAACCGCGCCAGACCGGCAAGGCCGGCTACGTG     2034
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTTCTACGTGGACGGCGCGCCCAACCGCGCCAGACCGGCAAGGCCGGCTACGTG     2040

Qy   2035 ACCGACCGGGGCGGCCAGAAGATCGTGAGCCTGACAGACCACCAGAGACCACCAGACGAGAACACCGAG     2094
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACCGGGGCGGCCAGAAGATCGTGAGCCTGACAGACCACCAGAGACCACCAGACGAGAACACCGAG     2100

Qy   2095 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC     2154
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC     2160

Qy   2155 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCAAGGCATCCAGCAAG     2214
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCAAGGCATCCAGCAAG     2220

Qy   2215 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC     2274
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC     2280

Qy   2275 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG     2334
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG     2340

Qy   2335 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGATTAAAAGCTTCCCGGGGCTAGCATTAGGACGACCTG     2394
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGATTAAAAGCTTCCCGGGGCTAGCATTAGGACGACCTG     2400

Qy   2395 TACGTGGGCAGCGGCGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCACCGGT     2451
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGGCAGCGGCGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCACCGGT     2457
```

RESULT 12
US-10-190-305A-39
; Sequence 39, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt_C
US-10-190-305A-39

Query Match   97.8%;   Score 2403.4;   DB 14;   Length 2457;
Best Local Similarity   99.3%;   Pred. No. 0;

RESULT 13
US-10-190-435-14
; Sequence 14, Application US/10190435
; Publication No. US2003014324A1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 3537
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagPolmutAtt_C
US-10-190-435-14

  Query Match   97.8%; Score 2402.8; DB 14; Length 3537;
  Best Local Similarity   99.1%; Pred. No. 0;
  Matches 2416; Conservative   0; Mismatches   22; Indels   0; Gaps   0;

```
Qy   374  TCCCCAGATCACCCTGTGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCA   433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1454  TCCCCAGATCACCCTGTGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCA  1513

Db   434  AGGAGGCCCTGCTGACGAGCACCGGCGGCGCGGACACCACCGGGATGAGCCTGCCCG   493
          |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Qy  1514  AGGAGGCCCTGCTGACGAGCACCGGCGGCGACGACCACCGGGATGAGCCTGCCCG   1573

Qy   494  GCAAGTGGAAGCCCAAGATCGGCGGCATCGGCGGCCTTCATCAAGGTGCGCCAGTACG   553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1574  GCAAGTGGAAGCCCAAGATCGGCGGCATCGGCGGCCTTCATCAAGGTGCGCCAGTACG  1633

Db   554  ACCAGATCCTGATCGAGATCGGCGGCAAGAGGCCATCGGCACCGTGTCGATCGGCCCCA   613
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1634  ACCAGATCCTGATCGAGATCGGCGGCAAGAGGCCATCGGCACCGTGTCGATCGGCCCCA  1693

Qy   614  CCCCCGTGAACATCATCGGCCGCCGCAGTGGGCTGGGCTGCACCCTGAACTTCC   673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1694  CCCCCGTGAACATCATCGGCCGCCGCAGTGGGCTGGGCTGCACCCTGAACTTCC   1753

Db   674  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGGCCCGGCATGGACGGCCCCAAGG   733
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1754  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGGCCCGGCATGGACGGCCCCAAGG  1813

Qy   734  TGAAGCAGTGGCCCCTGACCGAGAAGATCAAGGCCCTGACCGGCATCTGCGAGGAGA   793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1814  TGAAGCAGTGGCCCCTGACCGAGAAGATCAAGGCCCTGACCGGCATCTGCGAGGAGA  1873

Db   794  TGGGAGAAGGAGGGCAAGATCACCCAAGATCGGCCCCGAGAACCCCTACAACCCCGTGT   853
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1874  TGGGAGAAGGAGGGCAAGATCACCCAAGATCGGCCCCGAGAACCCCTACAACCCCGTGT  1933

Qy   854  TCGCCCATCAAGAAGCAGCAGCACCACCGGTGGCAGCTGCGCAAGCTGGTCGCAAGCTGCT   913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1934  TCGCCCATCAAGAAGCAGCAGCACCACCGGTGGCAGCTGCGCAAGCTGGTCGCAAGCTGCT  1993

Db   914  ACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGCATCCCCACCCCCGCGCCTGA   973
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1994  ACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGCATCCCCACCCCCGCGCCTGA   2053

Qy   974  AGAGAAGAGAGAGCCGTGACCGTGTCGCGACGCGTGGGCGACGTGGGGCCGTGCCCCTGG  1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2054  AGAGAAGAGAGAGCCGTGACCGTGTCGCGACGCGTGGGCGACGTGGGGCCGTGCCCCTGG  2113

Db  1034  ACGAGGACTTCGGCAAGTCACACCGCCATCCCCAACATCAACAACGAGACACCCCCG  1093
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2114  ACGAGGACTTCGGCAAGTCACACCGCCATCCCCAACATCAACAACGAGACACCCCCG  2173

Qy  1094  GCATCCGCTACCAGTACAAGTGTCTGCCGCCAGGGCTGAAGGGCAGCCCAGCATCTTC  1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2174  GCATCCGCTACCAGTACAAGTGTCTGCCGCCAGGGCTGAAGGGCAGCCCAGCATCTTC  2233

Db  1154  AGAGCAGCATGACGACCAAGATCCTTCGGCGCGCCTTCGGCGCCGTGACCATCGTGATCT  1213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2234  AGAGCAGCATGACGACCAAGATCCTTCGGCGCGCCTTCGGCGCCGTGACCATCGTGATCT  2293

Qy  1214  ACCAGGCCCCCTGTACGTGGGCGACCTGGAGATCGGCGACGCGGCCAGCAGCGGCCAAGATCG  1273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2294  ACCAGGCCCCCTGTACGTGGGCGACCTGGAGATCGGCGACGCGGCCAGCAGCGGCCAAGATCG  2353

Db  1274  ACCAGGCCCCCCGTGACCTGCAAGCACCGCTTCACCCCGACAAGAGAAGAGCACCAGA  1333
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2354  ACCAGGCCCCCCGTGACCTGCAAGCACCGCTTCACCCCGACAAGAGAAGAGCACCAGA  2413

Qy  1334  AGGAGGCTGGCAAGCTGCTGCTGCGCGCCCATCGAGCTGCACCCCGACAAGTCGGACGAGC CGAGCCATCG  1393
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1393  AGGAGGCTGGCAAGCTGCTGCTGCGCGCCCATCGAGCTGCACCCCGACAAGTCGGACGAGCCGAGCCATCG  2473

Qy  1394  AGCTGCCCCGAGAAGGAGAGAGCTGAACGACCGTGAACCGACATCCAGAGAGTTGGGGCAAGCTGA  1453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2474  AGCTGCCCCGAGAAGGAGAGAGCTGAACGACCGTGAACCGACATCCAGAGAGTTGGGGCAAGCTGA  2533
```

```
Qy  1454  ACTGGGCAGCAGCAGATCTACCCGGCGCATCAAGGTGCCGCCAGTGTGCAAGCTGCTGCGCG  1513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2534  ACTGGGCAGCAGCAGATCTACCCGGCGCATCAAGGTGCCGCCAGTGTGCAAGCTGCTGCGCG  2593

Qy  1514  GGGCCAAGGCCCTGACCGACATCGTGCCCGACGAGGAGCCCGAGCTGGACGCCCTGGAGCTGGCCG  1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2594  GGGCCAAGGCCCTGACCGACATCGTGCCCGACGAGGAGCCCGAGCTGGACGCCCTGGAGCTGGCCG  2653

Qy  1574  AGAACCGCAGATCCTGCGCGAGCCGCGCAGGGCGTGTACTACGACCCCAGCAAGGACC  1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2654  AGAACCGCAGATCCTGCGCGAGCCGCGCAGGGCGTGTACTACGACCCCAGCAAGGACC  2713

Qy  1634  TGGTGGCCGAGATCCAGAAGCCAGCAGGGCCCACGACCAGCAGTGACTACCAGCAGC  1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2714  TGGTGGCCGAGATCCAGAAGCCAGCAGGGCCCACGACCAGCAGTGACTACCAGGAGC  2773

Qy  1694  CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGACCGCCGCCACCAACGACG  1753
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2774  CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGACCGCCGCCACCAACGACG  2833

Qy  1754  TGAAGCAGTGACCGAGGCCGTGCAGAAGATGCCATGGAGAGACCTGGGACGACTGGGACCGACT  1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2834  TGAAGCAGTGACCGAGGCCGTGCAGAAGATGCCATGGAGAGACCTGGGACGACTGGGACCGACT  2893

Qy  1814  AGACCCCCAAGTTCGGCCTGCCCATCCAGAAGGAACACCCCCCCCTGGTGAAGC  1873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2894  AGACCCCCAAGTTCGGCCTGCCCATCCAGAAGGAACACCCCCCCCTGGTGAAGC  2953

Qy  1874  ACTGGCAGGCCACCTGAATCCCCGAGTTCGTGACTTCGTGAACAACCCCGAGAGC  1933
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2954  ACTGGCAGGCCACCTGAATCCCCGAGTTCGTGAACAACCCCGAGAGC  3013

Qy  1934  TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCCGCGGGGGCCCGGCCAGA  1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3014  TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCCGCGGGGGCCCGGCCAGA  3073

Qy  1994  CCGCCAACCGCGAGACCAAGATCGACGAGGCGCTGATCCAGCTCG  2053
          ||||||||||||||||||||||||||||||||||||||||||||
Db  3074  CCGCCAACCGCGAGACCAAGATCGACGAGGCGCTGATCCAGCTGG  3133

Qy  2054  AGATCGTGAGCTGGACGCCACACCAGAGACCCGAGCTCGCAGCCCTGGGCA  2113
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3134  AGATCGTGAGCTGGACGCCACACCAGAGACCCGAGCTCGCAGCCCTGGGCA  3193

Qy  2114  CCCTGCCAGGACAGCGGCAGCGGCCAGCAGGTGAACATCGTGACCGACAGCCCCTGGGCA  2173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3194  CCCTGCCAGGACAGCGGCAGCGGCCAGCAGGTGAACATCGTGACCGACAGCCCCTGGGCA  3253

Qy  2174  TCATCCAGGCCCAGCCCCAGCAGCGAGCTGGTGAACCAGAGCTGGTGCCCGCCGCA  2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3254  TCATCCAGGCCCAGCCCCAGCAGCGAGCTGGTGAACCAGAGCTGGTGCCCGCCGCA  3313

Qy  2234  TGATCAAGAAGGAGAAGGTGTACTGAGCTGGTGCCCGCCCCACAAGGGCATCGGCGGCA  2293
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3314  TGATCAAGAAGGAGAAGGTGTACTGAGCTGGTGCCCGCCCCACAAGGGCATCGGCGGCA  3373

Qy  2294  ACGAGCAGATCGACAAGTCGTGAGCAAGGACATCCGCAAGGTGCTGTTCCTGGGCGCA  2353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3374  ACGAGCAGATCGACAAGTCGTGAGCAAGGACATCCGCAAGGTGCTGTTCCTGGGCGCA  3433

Qy  2354  TCGATGGCGGCATCGTGATCTCACCAGTACATGGACAACGTGTACGGGCAGCGGCGGCC  2413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3434  TCGATGGCGGCATCGTGATCTCACCAGTACATGGACAACGTGTACGGGCAGCGGCGGCC  3493

Qy  2414  CTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT  2451
          |||||||||||||||||||||||||||||||||||||
Db  3494  CTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT  3531
```

; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 15
;   LENGTH: 3537
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagPolmutina_C
US-10-190-435-15

Query Match   97.8%; Score 2402.8; DB 14; Length 3537;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2416; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   14 TGGCCCGAGGCCATGAGCCAGGCCACCATCCTCGATGCAGCCCAGCAACTTCA 73
           ||||||||||||||||||||| |||||||||||||||||||||| |||||||||
Db   1094 TGGCCCGAGGCCATGAGCCAGGCCAACACCACCAGCGTGATGATGCAGAAGAGCAACTTTAAAA 1153

Qy   74 AGGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA 133
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1154 AGGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA 1213

Qy   134 ACTGCGGCGCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1214 ACTGCGGCGCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 1273

Qy   194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCTTCCCCCAGGGCA 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 1333

Qy   254 AGGCCCGCGAGTTCCCAGCAGCAGAACCGCGCCAACGCCCCAAGACCCGCGGAGCTGC 313
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   1334 AGGCCCGCGAGTTCCCCAGCAGCAGAACCGCCAACAGCCGGCACCCGGCGAGCTGC 1393

Qy   314 AGGTGCGCGGCGACAACCCCGCAGCCCCCGCGCCGCGAGCGCCCCGAGCGCCCAGGGCCAGCCCTGAACT 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1394 AGGTGCGCGGCGACAACCCCGCAGCCCCCGCGCCGCGAGCGCCCCGAGCGCCCAGGGCCAGCCCTGAACT 1453

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1454 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA 1513

Qy   434 AGGAGGCCCTGCTGGACACCGGCGCGGACACCGTGCTGGAGGAGATGAGCCTGCCCG 493
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514 AGGAGGCCCTGCTGGACACCGGCGCGGACACCGTGCTGGAGGAGATGAGCCTGCCCG 1573

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
           ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   1574 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCGGCTTCATCAAGGTGCGCCAGTACG 1633

Qy   554 ACCAGATCCTGATCGAGATCTGCGGCAAGAGCTGCCATCGGCACCGTGCTGCTGATCGGCCCCA 613
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1634 ACCAGATCCTGATCGAGATCTGCGGCAAGAGCTGCCATCGGCACCGTGCTGCTGATCGGCCCCA 1693

Qy   614 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGCTGCACCCTGAACTTCC 673
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1694 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGCTGCACCCTGAACTTCC 1753

Qy   674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 733
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1754 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 1813

Qy   1754 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 1813

Db   14 TGAAGCAGTGGCCCTGACCGAGGAGGAAGATCAGGCCCTGACCGCCATCTGGAGGAGA 793
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   734 TGAAGCAGTGGCCCTGACCGAGGAGGAAGATCAGGCCCTGACCGCCATCTGGAGGAGA 1814

Db   1814 TGGAGAAGGAGGCAAGATCACCAAGATGGCCCCGAGAACCCTACAACACCCCGTGT 853
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   794 TGGAGAAGGAGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGT 1874

Db   854 TCGCCATCAAGAAGAGGACAGCAGCCACCAAGGTGGTGGACTTCCGCGAGCTGA 913
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1934 TCGCCATCAAGAAGAGGAGGACAGCACCAAGGTGCGCGTGGACTTCCGCGAGCTGA 913

Db   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTTGGCCATCCCCACCCGCCGGCTGA 973
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1994 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTTGGGCCATCCCCACCCGCCGGCCTGA 2053

Db   974 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2054 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 2113

Db   1034 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGCACCCCG 1093
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2114 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGCACCCCG 2173

Db   1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCTGAAGGGCAGCCCAGCATCTCC 1153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2174 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCTGAAGGGCAGCCCAGCATCTCC 2233

Db   1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCAGGATCGTGATCT 1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2234 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCCGAGATCGTGATCT 2293

Db   1214 ACCAGGCCCCCTGTACGTGGGCGACCTGGAGATGGGCCAGCGCGCCAAGATCG 1273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2294 ACCAGGCCCCCTGTACGTGGGCGACCTGGAGATCGGGCCAGCGCGCCAAGATCG 2353

Db   1274 AGGAGCTGCGCAAGCACCTGCTGTCGCTGGGCTTCACCACCCCGACAAGAAGCACCAGA 1333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2354 AGGAGCTGCGCAAGCACCTGCTGTCGCTGGGCTTCACCACCCCGACAAGAAGCACCAGA 2413

Db   1334 AGGAGCCCCCTTCTGCCCATCGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCG 1393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2414 AGGAGCCCCCCTTCTGCCCATCGAGCTGCACCCCGACAGTGGACCGTGCAGCCCATCG 2473

Db   1394 AGTCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 1453
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Qy   2474 AGTCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 2533

Db   1454 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCG 1513
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2534 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCG 2593

Db   1514 GCGCCAAGGCCCTGACCGACGTGTCCCCTGGCACGGCGGTGTACTACGACCCCAGCAAGGACC 1573
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2594 GCGCCAAGGCCCTGACCGACGTGTCCCCTGACCGGAGGGCCGGTGTACTACGACCCCAGCAAGGACC 2653

Db   1574 AGAACCGGCAGATCTGCCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 1633
           ||||||| |||||||||||||||||||||||||||||||||||||||||||
Qy   2654 AGAACCGGCAGATCTGCCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 2713

Db   1634 TGGTGGCCGAGATCCAGAAGGCCGAGAAGGCCATCGGCCAAGTGACCGACCTACCAGGAGC 1693
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2714 TGGTGGCCGAGATCCAGAAGGCCGAGAAGGCCATCGGCCAAGTCTACCAGGAGC 2773

Db   1694 CCTTCAAGAACCTGAAGACCGGCCAAGTACGCCACCGCCAACGACG 1753
           |||||||||||||||||||||||||||||||||||||||||||||
Qy   2774 CCTTCAAGAACCTGAAGACCGGCCAAGTACGCCACCGCCAACGACG 2833

Db   1754 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATCGAGAGCATCGTGGGCCA 1813
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2834 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATCGAGAGCATCGTGGGGGCA 2893

Db   1873 AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCGACT 1873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1814 AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCGACT 1873

Db   14   TGGCCGAGGCCATGAGCCAGGCCACCAGCGCGCCAACATCCTGATGCAGCGCAGCAACTTCA   73
Qy 1487   TCGCCGAGGCCATGAGCCAGGCCACCAGCGCGCCAACATCCTGATGCAGCGCAGCAACTTCA   1546

Qy   74   AGGGCCCAAGGCCATCATCAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA   133
Db 1547   AGGGCCCAAGGCCATCATCAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA   1606

Qy   134   ACTGCCGCGGCCCCCCCGCAAGAAGGCGTGCTGGAAGTGCGGCAAGGTGGGCCACCAGATGA   193
Db 1607   ACTGCCGCGGCCCCCCCGCAAGAAGGCGTGCTGGAAGTGCGGCAAGGTGGGCCACCAGATGA   1666

Qy   194   AGGACTGACCAGGACCAACTTCTTCCGGCAAGGACCTGCCTTCCCCAGGGCA   253
Db 1667   AGGACTGACCAGGACCAACTTCTTCCGGCAAGGACCTGCCTTCCCCAGGGCA   1726

Qy   254   AGGCCGCGAGTTCCCAGCGAGCAGAACCGGCCAACAGCCCCACCAGCCGCGAGCTGC   313
Db 1727   AGGCCGCGAGTTCCCAGCGAGCAGAACCGGCCAACAGCCCCACCAGCCGCGAGCTGC   1786

Qy   314   AGGTGCGCGGCGACAACCCCGCGAGGCCCGAGCGCGGCCAGGGCACCTGGAACT   373
Db 1787   AGGTGCGCGGCGACAACCCCGCGAGGCCCGAGCGCGGCCAGGGCACCTGGAACT   1846

Qy   374   TCCCCCAGATCACCCTGTGGCAGCGCCCCCCTGGTGAGGATCAAGGTGGCGGCCAGATCA   433
Db 1847   TCCCCCAGATCACCCTGTGGCAGCGCCCCCCTGGTGAGGATCAAGGTGGCGGCCAGATCA   1906

Qy   434   AGGAGGGCCCTGCTGGACACCGGCGCCCGACGACACCGTGCTGGAGGATGAGCCTGCCCG   493
Db 1907   AGGAGGGCCCTGCTGGACACCGGCGCCCGACGACACCGTGCTGGAGGATGAGCCTGCCCG   1966

Qy   494   GCAAGTGGAAGCCCAAGATGATCGGCAAGATGATCTGCGCCGAACATGCTGACCAAGTGCGCCAGTACG   553
Db 1967   GCAAGTGGAAGCCCAAGATGATCGGCAAGATGATCTGCGCCGAACATGCTGACCAAGTGCGCCAGTACG   2026

Qy   554   ACCAGATCCTGATCGAGATCTGCGCCGAACAAGGCCATCGGCCACCGTGCTGATCGGCCCCA   613
Db 2027   ACCAGATCCTGATCGAGATCTGCGCCGAACAAGGCCATCGGCCACCGTGCTGATCGGCCCCA   2086

Qy   614   CCCCGTGAACATCATCGGCCGCCAACATGCTGACCAAGTGGCCGCAGCCCTGAACTTCC   673
Db 2087   CCCCGTGAACATCATCGGCCGCCAACATGCTGACCAAGTGGCCGCAGCCCTGAACTTCC   2146

Qy   674   CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG   733
Db 2147   CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG   2206

Qy   734   TGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA   793
Db 2207   TGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA   2266

Qy   794   TGGGAAGGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT   853
Db 2267   TGGGAAGGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT   2326

Qy   854   TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCCCAAGTGGTTGGACTTCCGGAGCTGA   913
Db 2327   TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCCCAAGTGGTTGGACTTCCGGAGCTGA   2386

Qy   914   ACAGGCGACCACCAGGACTTCTGGGAGGTGCAGTGGCATCCCCACCCGGCCTGA   973
Db 2387   ACAAGCGCACCACCAGGACTTCTGGGAGGTGCAGTGGCATCCCCACCCGGCCTGA   2446

Qy   974   AGAAGAAGAGAGGGTGCGACGGTGGGCGACCGCTACTTCAGCGTGCCCCTGG   1033
Db 2447   AGAAGAAGAGAGGGTGCGACGGTGGGCGACCGCTACTTCAGCGTGCCCCTGG   2506

Qy 1034   ACGAGGGACTTCCGGCAAGTACACCCGGCCTTCACCATCCCCAGCATCAACAACACGAGACCCCCG   1093
Db 2507   ACGAGGGACTTCCGGCAAGTACACCCGGCCTTCACCATCCCCAGCATCAACAACACGAGACCCCCG   2566

Db 2894   AGACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGGACCTGGGAGACCTGGTGGACCGGACT   2953

Qy 1874   ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGGTGAAGC   1933
Db 2954   ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGGTGAAGC   3013

Qy 1934   TGTGGTACCAGCTGGAGAAGGAGGCCATCATCGGGCGCGCAGACCTTCTACGTGACGGGCG   1993
Db 3014   TGTGGTACCAGCTGGAGAAGGAGGCCATCATCGGGCGCGCAGACCTTCTACGTGACGGGCG   3073

Qy 1994   CCGCCAACCGGGAGACCCGGAGACCAAGATGGGCAAGGCGGCTACGTGACCAGCGGGCCGCAGA   2053
Db 3074   CCGCCAACCGGGAGACCCGGAGACCAAGATGGGCAAGGCGGCTACGTGACCAGCGGGCCGCAGA   3133

Qy 2054   AGATCGTGAGCGTGACCGAGACCACCAACCAGAAGACCAGCTGCAGGCAGGCCATCCAGGTCG   2113
Db 3134   AGATCGTGAGCGTGACCGAGACCACCAACCAGAAGACCAGCTGCAGGCAGGCCATCCAGGTGG   3193

Qy 2114   CCCTGCAGGACAGCGGCGGCCAGCGAGTGAACATCGTGACCGACACCGTGCCCGACTACGCCGTCTGGGCA   2173
Db 3194   CCCTGCAGGACAGCGGCGGCCAGCGAGTGAACATCGTGACCGACACCGTGCCCGACTACGCCGTCTGGGCA   3253

Qy 2174   TCATCCAGGCCCCAGCCCGACAAGAGCCGAGAGCGAGGCCGAGCTGGTGAACCAGATCATCGAGACCAGC   2233
Db 3254   TCATCCAGGCCCCAGCCCGACAAGAGCCGAGAGCGAGGCCGAGCTGGTGAACCAGATCATCGAGACCAGC   3313

Qy 2234   TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA   2293
Db 3314   TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA   3373

Qy 2294   ACGAGCAGATCGACAAGCTGGTGAGCAGCGGCATCCGCAAGGTGCTGTTCCTGGACGGCA   2353
Db 3374   ACGAGCAGATCGACAAGCTGGTGAGCAGCGGCATCCGCAAGGTGCTGTTCCTGGACGGCA   3433

Qy 2354   TCGATGGCGGCATCGTGATCTACCAGTACTATGGACGACGACCTTGTACGTGGGCAGCGCGCGGCC   2413
Db 3434   TCGATGGCGGCATCGTGATCTACCAGTACATGACACATGGACGACGACCTTGTACGTGGGCAGCGCGCGGCC   3493

Qy 2414   CTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT   2451
Db 3494   CTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT   3531

Query Match   97.8%; Score 2401.8; DB 14; Length 5145;
Best Local Similarity   99.9%; Pred. No. 0;
Matches 2403; Conservative   0; Mismatches   0; Indels   2; Gaps   0;

```
Qy   1094  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC  1153
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   2567  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC  2626

Qy   1154  AGAGCAGCATGACCAAGATCCTGGAGACCCTTCGCGGCACCCCGAGATCGTGATCT      1213
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   2627  AGAGCAGCATGACCAAGATCCTGGAGACCCTTCGCGGCACCCCGAGATCGTGATCT      2686

Qy   1214  ACCAGGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCAGCACCGCGCCAAGATCG     1273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2687  ACCAGGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCAGCACCGCGCCAAGATCG     2746

Qy   1274  AGGAGCTGCGCAAGCACCTGCTGCCGCTGGGCGGCTTCACCACCCCGACAAGAAGCACCAGA  1333
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2747  AGGAGCTGCGCAAGCACCTGCTGCCGCTGGGCGGCTTCACCACCCCGACAAGAAGCACCAGA  2806

Qy   1334  AGGAGCCCCCTTCCTGCCCATCGAGCTGCACCCCGACAAGCGTGACCGTGCAGCCCATCG  1393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2807  AGGAGCCCCCTTCCTGCCCATCGAGCTGCACCCCGACAAGCGTGACCGTGCAGCCCATCG  2866

Qy   1394  AGTCGCCCGAGAAGGAGAGTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA   1453
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2867  AGTCGCCCGAGAAGGAGAGTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA   2926

Qy   1454  ACTGGGCCAGCCAGATCGTACCCCGGCGCATCGTGCCCCTGACCGACATCGTGCCCGAGCTGCTGCGCCG  1513
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2927  ACTGGGCCAGCCAGATCTACCCCGGCGCATCGTGCCCCTGACCGACATCGTGCCCGAGCTGCTGCGCCG  2986

Qy   1514  GCGCCAAGGCCCTGACCGACATCGTGCCCGAGCCGTGCACGGCCGTGTACTACGACCCCAGCAAGGACC  1573
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2987  GCGCCAAGGCCCTGACCGACATCGTGCCCGAGCCGTGCACGGCCGTGTACTACGACCCCAGCAAGGACC  3046

Qy   1574  AGAACCGCGAGATCCTGCGGAGATCCTGCGCACGGCCGTGTACTACGACCCCAGCAAGGACC  1633
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3047  AGAACCCGCGAGATCCTGCGGAGATCCTGCGCACGGCCGTGTACTACGACCCCAGCAAGGACC  3106

Qy   1634  TGGTGGCCGAGATCCAGAGATCCAGGAGGCAGGGGCCACGACCAGTGGACCTTACCAGGAGC  1693
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3107  TGGTGGCCGAGATCCAGAGATCCAGGAGGCAGGGGCCACGACCAGTGGACCTTACCAGGAGC  3166

Qy   1694  CCTTCAAGAACCTGAAGACCGGCAAGATGCGCAAGATGCCAAGATCGGCCACCACCACGACG  1753
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3167  CCTTCAAGAACCTGAAGACCGGCAAGATGCGCAAGATGCCAAGATCGGCCACCACCACGACG  3226

Qy   1754  TGAAGCAGCTGACCGAGCCGTGCAGAAGATCGCCATGGAGAGAGATCCCATGGAGAGCATCGTGATCTGATCTGGGGCA  1813
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3227  TGAAGCAGCTGACCGAGCCGTGCAGAAGATCGCCATGGAGAGAGATCCCATGGAGAGCATCGTGATCTGATCTGGGGCA  3286

Qy   1814  AGACCCCCAAGTTCGCCTGCCTCATCGAAGGAGAACCTGGGAGACCTGGTGACCGACT   1873
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3287  AGACCCCCAAGTTCGCCTGCCTCATCGAAGGAGAACCTGGGAGACCTGGTGACCGACT   3346

Qy   1874  ACTGGCAGGCCACCTGGATCCCCGAGTTCGTGAAACACCCCCCCCCCTGGTGAAGC   1933
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3347  ACTGGCAGGCCACCTGGATCCCCGAGTTCGTGAAACACCCCCCCCCCTGGTGAAGC   3406

Qy   1934  TGTGGTACCAGCTGGAGAAGGAGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCG  1993
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3407  TGTGGTACCAGCTGGAGAAGGAGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCG  3466

Qy   1994  CCGCCAACCGCGAGACCAGATCGGCCAAGGCCGCCGCTACGTGACCGGGGCCGGCAGA  2053
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3467  CCGCCAACCGCGAGACCAGATCGGCCAAGGCCGCCGCTACGTGACCGGGGCCGGCAGA  3526

Qy   2054  AGATCGTGAGCGCCTGACCGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG  2113
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3527  AGATCGTGAGCGCCTGACCGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG  3586

Qy   2114  CCCTTGCAGGACACGGCGGCAGCCAGCGAGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCG  2173
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3587  CCCTTGCAGGACACGGCGGCAGCCAGCGAGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCG  3646

Qy   2174  TCATCCAGGCCCAGCCCGACAAGAGACCCGAGAGGCGGAGCGAGCTGGTGAACCAGATCATCGAGCAGC  2233
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3647  TCATCCAGGCCCAGCCCGACAAGAGACCCGAGAGGCGGAGCGAGCTGGTGAACCAGATCATCGAGCAGC  3706

Qy   2234  TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTCCCGCCCACAAGGGCATCGGCGCGGCA  2293
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3707  TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGCCGCCCACAAGGGCATCGGCGCGGCA  3766

Qy   2294  ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCCAAGGTGCTGTTCCTGGACGGCA  2353
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   3767  ACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCCAAGGTGCTGTTCCTGGACGGCA  3826

Qy   2354  TCGATGGCGGCATCGTGATCTACCAGTACATGACGACGACCTGTACGTGGGCAGCGGCGGCC  2413
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3827  TCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCGGCC    3886

Qy   2414  CTAGG  2418
           |||||
Db   3887  CTAGG  3891
```

Search completed: April 10, 2004, 21:24:47
Job time : 597.888 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: April 10, 2004, 06:48:32 ; Search time 118.045 Seconds
(without alignments)
11550.795 Million cell updates/sec

Title: US-09-610-313-32
Perfect score: 2457
Sequence: 1 gtcgagcgcaccatgccga........gggctagcaccggtgaattc 2457

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2//ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2//ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2//ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2//ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2//ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2//ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	2040	83.0	2300	4	US-09-475-515-83	Sequence 83, Appl
2	2019.2	82.2	2306	4	US-09-475-515-82	Sequence 82, Appl
3	1998.4	81.3	2312	4	US-09-475-515-84	Sequence 84, Appl
4	1915.8	78.0	4319	4	US-09-475-515-6	Sequence 6, Appli
5	1872.8	76.2	2299	4	US-09-475-515-81	Sequence 81, Appl
6	1852	75.4	2305	4	US-09-475-515-80	Sequence 80, Appl
7	1626.6	66.2	4307	4	US-09-552-950-2	Sequence 2, Appli
8	1599.4	65.1	9772	4	US-09-552-950-5	Sequence 5, Appli
9	1539.2	62.6	8366	4	US-09-872-733A-6	Sequence 6, Appli
10	1503.8	61.2	4338	4	US-09-872-733A-1	Sequence 1, Appli
11	1178.4	48.0	9010	4	US-09-184-418C-8	Sequence 8, Appli
12	1147.4	46.7	8972	4	US-09-184-418C-9	Sequence 9, Appli
13	1140.2	46.4	8959	4	US-09-184-418C-11	Sequence 11, Appli
14	1137	46.3	2467	4	US-09-872-733A-3	Sequence 3, Appli
15	1116.8	45.5	8992	4	US-09-184-418C-4	Sequence 4, Appli
16	1081.6	44.0	2601	3	US-09-117-217-7	Sequence 7, Appli
17	1081.6	44.0	2601	3	US-09-117-217-9	Sequence 9, Appli
18	1081.6	44.0	2601	3	US-09-117-217-11	Sequence 11, Appli
19	1081.6	44.0	2601	3	US-09-117-217-13	Sequence 13, Appli
20	1081.6	44.0	2601	4	US-09-735-487-7	Sequence 7, Appli
21	1081.6	44.0	2601	4	US-09-735-487-9	Sequence 9, Appli
22	1081.6	44.0	2601	4	US-09-735-487-11	Sequence 11, Appli
23	1081.6	44.0	2601	4	US-09-735-487-13	Sequence 13, Appli
24	1081.6	44.0	4307	4	US-09-552-950-1	Sequence 1, Appli
25	1081.6	44.0	9719	4	US-09-700-304-1	Sequence 1, Appli
26	1080	44.0	9050	4	US-09-184-418C-7	Sequence 7, Appli
27	1075.2	43.8	7399	2	US-08-418-848A-9	Sequence 9, App

28	1075.2	43.8	9709	2	US-08-188-583-5	Sequence 5, Appli
29	1075.2	43.8	9709	3	US-08-388-353-1	Sequence 1, Appli
30	1075.2	43.8	9709	3	US-08-488-551B-1	Sequence 1, Appli
31	1075.2	43.8	9709	4	US-09-309-572-15	Sequence 15, Appl
32	1075.2	43.8	9709	4	US-09-718-096-15	Sequence 15, Appl
33	1075.2	43.8	12479	4	US-09-318-138-13	Sequence 13, Appl
34	1075.2	43.8	12494	3	US-08-935-312-13	Sequence 13, Appl
35	1075.2	43.8	12494	3	US-08-848-760B-33	Sequence 33, Appl
36	1075.2	43.8	15581	3	US-08-646-538-35	Sequence 35, Appl
37	1075.2	43.8	15581	3	US-09-503-222-35	Sequence 35, Appl
38	1074	43.7	3000	4	US-09-184-418C-74	Sequence 74, Appl
39	1070.6	43.6	8968	4	US-09-184-418C-1	Sequence 1, Appli
40	1065.6	43.4	4313	4	US-09-475-515-3	Sequence 3, Appli
41	1065.6	43.4	9737	2	US-08-944-449-7	Sequence 7, Appli
42	1065.6	43.4	9737	4	US-09-353-362-7	Sequence 7, Appli
43	1057.6	43.0	8954	4	US-09-184-418C-6	Sequence 6, Appli
44	1052.4	42.8	3017	4	US-09-184-418C-83	Sequence 83, Appl
45	1050.2	42.7	3011	4	US-09-184-418C-101	Sequence 101, App

ALIGNMENTS

RESULT 1
US-09-475-515-83
; Sequence 83, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(-).protmod.RTopt.YMWM
US-09-475-515-83

Query Match   83.0%; Score 2040; DB 4; Length 2300;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2153; Conservative 0; Mismatches 135; Indels 12; Gaps 2;

Qy   170 GCGCAAGGAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCAGGCCAACTTCTTCCGCG 229
             |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   1 GCGGCCCGCGGAGGACCACCAAATGAAGATTGCACTGAGAGACAGGCTAATTTCTTCCGCG 60

Qy   230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCAGCGAGGAGAACCGCGCGCCA 289
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db   61 AGGACCTGGCCTTCTGCGGGCAAGGCCCGCGAGTTCAGCAGCGAGGAGACCCGCGCGCCA 120

Qy   290 ACAGCCCACCACCAGCCGCGAGCTGCAGTGCGCGGCG------ACAACCCCGCGCAGCGGAGG 343
             |||||||||||||||||||||||||||||||||||| ||||||| ||||||||
Db   121 ACAGCCCCACCACCCGCGGCCGCGAGCTGCAGTGTGGGGGCGGCGGCGAGAACAACAGCCTGAGCGGAGG 180

Qy   344 CCGGCGCCGAGCGCGAGCGCCAGGGCACCCTG------AACTTCCCCCAGATCACCCTGTGGCAGC 397
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db   181 CCGGCGCCGACCGCCGCGGCCACCGTGAGCTTCAACTTCCCCCAGATCACCCCTGTGGCAGC 240

```
Qy    398  GCCCCCTGTGTGACGATCAAGGTGGGGCGGCCAGATCAAGGAGGCCCTGCTGACACCGGCG  457
           ||||||||||||||||| ||| || ||||| |||||| |||||||| |||||||||||||||
Db    241  GCCCCCTGTGTGACCATCGAGGATCGGCGGCCAGCTCAAGGAGCGCGTGCTGCTGACACCGGCG  300

Qy    458  CCGACGACACCGGTGCTGCTGGAGGAGATGAGCGGCAAGTGAAGCCCAAGATGATCG  517
           |||||||||||  |||||||||||  |||||||||||||||||||||||||||||||
Db    301  CCGACGACACCGGTGCTGCTGGAGGAGATGAAGTGGAAGCCCAAGATGATCG  360

Qy    518  GCGGCATCGGCGGCTTCATCAAGGTGCGCAGTACGACCAGATCCTGATCGAGATTGCG  577
           ||||||||||||||||||||||||  ||||||||||| ||||||||||||||||||||
Db    361  GCGGGATCGGCGGCTTCATCAAGGTGCGCAGTACGACCAGATCCTGGAGATCTGCG  420

Qy    578  GCAAGAAGGCCATCGGCGACCGTGCTGATCGGCCCCACCCCGTGAAACATCGGCCGCA  637
           ||| |||||||||||| |||||  ||||||||||||||||||||||  |||||||||||
Db    421  GCCACAAGGCCATCGGCGATCGGTGGGTCGGCCCCACCCCGTGAACATCATCGGCCGCA  480

Qy    638  ACATGCTGACCCAGCTGGGCTGCCACCCTGAACTTCCCATCAGCCCCATGAGACCGTGC  697
           ||| ||||||||||  ||||||||||||||| ||||||||||||||||  ||||||||||
Db    481  ACCTTGCTGACCCAGATCGGCTGCCACCCTGAACTTCCCATCAGCCCCATGAGACGCGTGC  540

Qy    698  CCGTGAAGCTGAAGCCCGGCATGACGACGCCCTGCGAGGTGAAGCAGTGCCCCAAGATCACCA  757
           ||||||||||||||||||||||||  |||||||||||||||| ||||||||| |||||||||||
Db    541  CCGTGAAGCTGAAGCCCGGGATGACGGGCCCTGCGAGGTGAAGCAGTGCCCCAAGATCAGCA  600

Qy    758  AGAAGATCAAGGCCCTGAGAACCCTACAAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCA  817
           ||||||||||||| |||  |||||||||||||||||||||||| |||||||||||||||||||||
Db    601  AGAAGATCAAGGCCCTGGAGATGTGGAGATCTGACCCGAGATCTGGAGAAGAAGGACAGCA  660

Qy    818  AGATCGGCCCCGAGAACCCTACAAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCA  877
           ||||| ||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Db    661  AGATCGGCCCCGAGAACCCTACAAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCA  720

Qy    878  CCAAGTGGCGCAAGCTGGTGGACTTCCGGCGAGCTGAACAAGCCGCACCCAGGACTTCTGGG  937
           ||||||||  |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    721  CCAAGTGGCGCAAGCTGGTGGACTTCCGGCGAGCTGAACAAGCCGCACCCAGGACTTCTGGG  780

Qy    938  AGGTGCAGCTGGGCGCATCCCCAGCATCAACAACAACAACAACGAGAGCGTGTACCGTTAC  997
           ||||||||||||||||||| ||||| |||||||||||||||||||||||||||||||||||
Db    781  AGGTGCAGCTGGGCGCATCCCCAGCATCAACAACAACAACAACGAGAGCGTGTACCGTTAC  840

Qy    998  TGGACGTGGGGCGACGCCTACTTCAGCGCGTGCCCCTGGAGCGGACTTCCGGACCTACACCG  1057
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TGGACGTGGGGCGACGCCTACTTCAGCGCGTGCCCCTGGAGCGGACTTCCGGACCTACACCG  900

Qy    1058 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGGCTACCAGTACAACGTGC  1117
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGGCTACCAGTACAACGTGC  960

Qy    1118 TGCCCCAGGGCTGGAAGGCAGCCCCAGCATCTTCAGGCAGCAGCCAAGATCCTGG  1177
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TGCCCCAGGGCTGGAAGGCAGCCCCAGCATCTTCAGGCAGCAGCCAAGATCCTGG  1020

Qy    1178 AGCCCTTCCGGCGCCCGCAACCCCGAGATCGTGATCTTACCAGGCCCCCGTACGTGGGGCA  1237
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    1021 AGCCCTTCCGGCGCCCGCAACCCCGAGATCGTGTAGATCGTGTACGTGGGGCA  1080

Qy    1238 GCGACCTGGAGATCGGCCAGCAACGACCACCGCGCCGACCAAGAAGAGGAGCTGCGCCAGGA  1297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 GCGACCTGGAGATCGGCCAGCAACGACCACCGCGCCGACCAAGAAGAGGAGCTGCGCCAGGA  1140

Qy    1298 GCTGGGGCTTCACCACCCCGACAAGAAGCACCACCCCCCCGACAAGAAGAGGAGCTGCCCATCG  1357
           |||||||||||||||||||||  |||||||||||||||||||||||||||||||||  |||||||
Db    1141 GCTGGGGCTTCACCACCCCGACAAGAAGCACCACCCCCCCGACAAGAAGAGGAGCTGCCCATCG  1200

Qy    1358 AGTGCACCCCGACAAGTGGACCGTGCCCAACCCCGAGACTTGCCCAGAAGGAGAGCTGGA  1417
           ||||||||||||||||| ||||||||||||||||||||| ||||||||||||||||||||
Db    1201 AGTGCACCCCGACAAGTGGACCGTGCCCAACCCCGAGACTTGCCCAGAAGGAGAGCTGGA  1260

Qy    1418 CCGTGAACGACATCCAGCTGGGCAAGCTGGGAACTGGGCCAGCCAGATCTACCCCG  1477
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CCGTGAACGACATCCAGCTGGGCAAGCTGGGAACTGGGCCAGCCAGATCTACGCCCG  1320

Qy    1478 GCATCAAGGTCGTGCTGCAAGCTGCTGTGCTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACACATCG  1537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1321 GCATCAAGGTGAAGCAGTCAGCTGTGCTGCGCGGCACCAGGCCCTGACCGACGGTGA  1380
           ||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Qy    1538 TGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGAGAACCTGCGCGGCGGCACCGAGAAGATCTGCGCGGAGC  1597
           |||||||||||||||||||  |||||||||||||||||||||||  |||||||||||||||||||||||||||
Db    1381 TCCCCCTGACCGAGGAGGAGCCGAGCTGGAGCTGGAGAGATCCTGAGAGAGC  1440

Qy    1598 CCGTGCACGGCGTGTTACTACGACCTGTGGCCAGATCCAGAGAAGCAGG  1657
           ||||||||       ||| ||| ||  |||||||||||||||||||||||
Db    1441 CCGTGCAGAGTGTACTACGACCTGGTGGCCCAGAGATCCAGAAGCAGG  1500

Qy    1658 GCCACGACCAGTGGACCACCAGATCTACCAGGAGAGCCCTCAAGAACCTGAAGAAGACCGGCA  1717
           ||| ||||||||||  ||||||||| ||||||||||| ||||||||||  |||||||||||||
Db    1501 GCCAGGGGCCAGTGGACCCTACCAGGATCTACCAGGAGAGCCTTCAAGAAACCTGAAGAACCGGCA  1560

Qy    1718 AGTACGCCAAGATGCCACCGCACCCGCCCCCAACGACGTGAAGCAGCTGACCGAGCTGCTGC  1777
           ||||  ||||| ||  |||||||||  ||||||||||||| ||||||  ||||||||||  ||
Db    1561 AGTACGCCCCGCATGCGGCGGGCCCCAACGACGTGAAGCAAGTGCACCCCAAGTTCCCGCC  1620

Qy    1778 AGAAGATCGCTACGAGAGCATCGTGATCTGTGGGACCGACTACTGGAGGCCACCTGA  1837
           ||| ||||||| ||||||||||||||||||||||||||||||| ||  ||||||||||
Db    1621 AGAAGGTGAGCACCGAGAGCATCGTGATCGGGTGGGATGACCTTGGTGACTGGAGGCCCA  1680

Qy    1838 TCCAGAGGAGACCTGGAGACCTGCAGGCCAAGATCACTGACACTCCCG  1897
           ||||||||||||| |||||||||||||| |||||||||||| |||||||
Db    1681 TCCAGAGGAGACCCTGGGAGGCCTGCAGGCCAAGGATCACTGGATCCCCG  1740

Qy    1898 AGTGGGAGTTCGTGAACACCCCCGTGTGTGAAGCTGTGTACAGCTGTGGTACGAGGAGC  1957
           |||  |||||||||| ||||||||||||||||| |||  |||||||| |||||||||||
Db    1741 AGTGGGAGTTCGTGAACAAGCTGTCCCCCGGTGGTGTGAAGCTGTGTTACCAGCTGGTGTACCAGAGGAGC  1800

Qy    1958 CCATCATCGGCGGCCGGCCGAGACCTTCTACGTGCACGGCAGCCCGAGAAGATCGTGAGCGCTGG  2017
           |||||||  ||  ||||| |||||||| ||  |||||| ||||  |||||||||||||||||||||
Db    1801 CCATCGTCGGGGGCCGGCCGAGACCTTCTACGTGGACGGCCGCCGGCCCAAGCCCAAGCTGTGG  1860

Qy    2018 GCAAGGCCGGCTACGTGACTGACCAGCCAGTGGAGGCCGTCAACAAGGCCGTGACCACCA  2077
           ||||||||||||||||||||||||||||||||||||  |||||||||  ||||||||||||
Db    1861 GCAAGGCCGGCTACGTGACTGACCAGCCAGTGGAGGCCGTGACATCGCCGGCGACAAGA  1920

Qy    2078 CCAACCAAGAAGACCGAGCTGAGCAGCTGCAGCCAGATATCGAGGAAGAGAGAGCAGCAGCGAGG  2137
           ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    1921 CCAACCAAGAAGACCGAGCTGAGCAGCTGCAGCCAGATCATCCACCTGCCCTGCGGCCTGGGAGG  1980

Qy    2138 TGAACATGTGACCGACCAGCCACCAGTTACCCGGCTACCAGTACAACGTGC  2197
           ||||||||||||||||||||||||  ||| |||||||||||||||||||||
Db    1981 TGAACATGTGACCGACCAGCCAGTGGCATCAGAGTTCCTGTTCCTGAACGGGCA  2040

Qy    2198 GCGAGAGGCCAGCTGGTGAACCAGATATCGAGCAGCTGGTGAACGAACGAGCAAGAAGGTGTACC  2257
           ||||||||||||||||||||||||||| |||||||||||||||  |||||||||||||||||||||
Db    2041 GCGGAGAGCCAGCTGGTGGAGCCAGATCATCGAGCAGCTGGTGAACAAGGGCATCAAGGAAGGTGTACC  2100

Qy    2258 TGAGCTGGTGGTGCCCCGCCAACAGGGGCAATCGTGGTGA  2317
           |||||||||||||||||||||||  |||||||||||||||
Db    2101 TGGCCTGGGTGCCCTGGGTGCCCCAAGGGCAGATGTGGACAAGCAGCAGTCGTGGTGA  2160

Qy    2318 GCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGTGCTGCGGCCATCGTCGTGATCTACC  2377
           |||||||||||||||||||||||||||||||||  |||||||| |||||||| |||||||
Db    2161 GCGCCGGCATCCGCAAGGTGCTGTTCCTGGAACGGCAGCCAAGGTGCTGTTCCTGGAGTGCTGTTCCTGGAAGTCGTGATCTACC  2220

Qy    2378 AGTACATGACGACCGTGTACGTGGGCAGCGGCGGGCCTAGGATGACTTCCCG  2437
           ||||||||||||||||||||||||||||||||||||||  |||||||||||||
Db    2221 AGTACATGACGACCGTGTACGTGGGCAGCGGCGGGCCCTAGGATGACTTCCCG  2280

Qy    2438 GGGCTAGCACCGTGAATTC  2457
           |||||||||||||||||||
Db    2281 GGGCTAGCACCGTGAATTC  2300
```

RESULT 2
US-09-475-515-82
; Sequence 82, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:

APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 82
LENGTH: 2306
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: FS(-).protmod.RTopt.YM

US-09-475-515-82

Query Match   82.2%;   Score 2019.2;   DB 4;   Length 2306;
Best Local Similarity   93.2%;   Pred. No. 0;
Matches 2150; Conservative   0; Mismatches 138; Indels 18; Gaps   3;

Db 1741 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         TCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA 1800

Qy 1952 AGGAGCCCATCATCGGCGCGAGACCTTCACGTGACGGCGCCCAACCGCGAGACCA 2011
         ||||||| |||||||||||||||||| ||||||||||| ||||||| |||||||||
Db 1801 AGGAGCCCATCGTGGGCGCCGAGACCTTCACGTGACGGCGCCCAACCGCGAGACCA 1860

Qy 2012 AGATCGGCAAGGCCGGCTACGTGACCGACCGCGGGCCGGCAGAAGATCGTGACCTGACCG 2071
         ||||||||| ||||||||||||||||||||||| |||||||||||||||||
Db 1861 AGCTCGGCAAGGCCGGCTACGTGACCGACCGCGGCGCGGCAGAAGGTGGTGAGCATCGCCG 1920

Qy 2072 AGACCACCAACCAGAGAGACCAGCTGCAGGCCATCCAGCTGCCCTGCAGGACAGCGGCA 2131
         ||||||||||||||||||||||||||||| |||||||||||||||||||
Db 1921 ACACCACCAACCAGAGAGACCAGCTGCAGGCCATCCACCTGGCCCTGCAGGACAGCGGCC 1980

Qy 2132 GCGAGGTGAACATCGTGACGGACCAGCAGCCAGTACGCCCTGGGCATCATCCAGCCCCG 2191
         |||||||||||||||||||||||||||||||||
Db 1981 TGGAGGTGAACATCGTGACCGACCAGCCAGTACGCCCTGGGCATCATCCAGCCCCGCCCCG 2040

Qy 2192 ACAAGAGCGAGAGCGAGCTGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG 2251
         ||||||||||||||||||||| |||||||||||||||||||||||||
Db 2041 ACAAGAGCGAGAGCGAGCTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG 2100

Qy 2252 TGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACCAAGC 2311
         ||||||| |||||||||||||||||||||
Db 2101 TGTACCTGGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGTGGACAAGC 2160

Qy 2312 TGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGGCATCGATGGCCGGCATCGTGA 2371
         |||||| ||||||||||||||||||||||||||||||||||
Db 2161 TGGTGAGCCGCCGCATCCGCAAGGTGCTGTTCCTGAACGGCATCGATGGCCGGCATCGTGA 2220

Qy 2372 TCTACCAGTACATGGACGACCTGTACGTGGGCAGCGCCCTAGGATCGATTAAAAGC 2431
         ||||||| ||||||||||||||||||||| ||||||||
Db 2221 TCTACCAGTACATGGACCACCTGTACGTGGGCAGCGCCCTAGGATCGATTAAAAGC 2280

Qy 2432 TTCCCGGGGCTAGCACCGGTGAATTC 2457
         ||||||||||||||||||||||||||
Db 2281 TTCCCGGGGCTAGCACCGGTGAATTC 2306

Qy 170 GCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGCGCCAGCCAACTTCTTCCGCG 229
         ||||| |||||||||||||||||||| || |||||||||||
Db 1   GCGGCGCCGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTCTTCCGCG 60

Qy 230 AGGAACCTGGCCTTCCCCAGGGCAAGGCCCGGAGTTCCCCAGCGGAGCGAGAACCGCGCCA 289
         ||||| |||||||||||||||||||||||||||
Db 61 AGGAACTGGCCTTCCTGCAGGGCAACCGGGCCGGGAGTTCAGCAGCGAGATCGTGACCGCCA 120

Qy 290 ACAGCCCACCAGCCGCGAGCTGCAGGTGCGCGCG-----ACAACCCCGCAGCGAGG 343
         ||||||||||||||||||||||||||| |||||
Db 121 ACAGCCCACCCGCCGCGAGCTGCAGCTGCGGCGGCGGGAGAACAACGACGCCTGAGCGAGG 180

Qy 344 CCGGCGCCGAGCGCCAGGGCACCCTG-----AACTTCCCCAGATCACCCTGTGGCAGC 397
         |||||||||||||||||||||||||
Db 181 CCGGGCGCCGACCGCCAGGGCACCCTGAGCTTCAACTTCCCCAGATCACCCTGTGGCAGC 240

Qy 398 GCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCAAGGAGCCCTGCTGGACACCGGCG 457
         ||||| |||||||||||||||
Db 241 GCCCCTGGTGACCATCAGGATCGGCGGCGAGCTCAAGGAGGGCGCTGCTGGACACCGGCG 300

Qy 458 CCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCG 517
         |||||||||||||||||||||||| |||||||||||||||
Db 301 CCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCG 360

Qy 518 GCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGCAGCCTGAGATCTGGG 577
         |||||||||||| |||||||||||||||
Db 361 GCGGGATCGGCGGGCTTCATCAAGGTGCGCCAGTACGACCAGCAGCCTGAGATCTGCG 420

Qy 578 GCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCGTGAACATCATCGGCCGCA 637
         |||| ||||||||||||||||||||||||||
Db 421 GCCACAAGGCCATCGGCACCGTGCTGATCGGCCCCGTGAACATCATCGGCCGCA 480

Qy 638 ACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCAGCCCGTGC 697
         |||||||||||||||||||||||||||||||||||
Db 481 ACCTGCTGACCCAGATCGGCTGCACCCTGAACTTCCCCATCAGCCCCATCAGCGACGGTGC 540

Qy 698 CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG 757
         ||||||||||||||||||||||||||||
Db 541 CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG 600

Qy 758 AGAAGATCAAGGCCCTGACCGCCATCTGGGAGGAGGATGGAAGGAGGCAAGATCACCA 817
         ||||||||||||||||||| |||| |||||||||||||||
Db 601 AGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGGAGGGCAAGATCAGCA 660

Qy 818 AGATCGGCCCCGAGAACCCCTACACACACCCCGTGTTCGCCATCAAGAAGAAGACAGCA 877
         ||||||||||||||||||||||||||||||||||||
Db 661 AGATCGGCCCCGAGAACCCCTACACACACCCCGTGTTCGCCATCAAGAAGAAGACAGCA 720

Qy 878 CCAAGTGGCGCAAGCTGGTGACTTCGGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 937
         |||||||||||||||||||||||||||||||||||||||
Db 721 CCAAGTGGCGCAAGCTGGTGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 780

Qy 938 AGGTGCAGCTGGGCATCCCCACCCCGCCGGCCTGAAGAGAGAGAGCGTGACCGTGC 997
         |||||||||||||||||| |||||||||||||||||||
Db 781 AGGTGCAGCTGGGCATCCCCACCCCGCCGGCCTGAAGAGAGAGAGCGTGACCGTGC 840

Qy 998 TGGACGTGGGCGACGCCTACTTCAGGTGCCCCTGACCGAGTACACCG 1057
         |||||||||||||||||||||||||||||||||| ||||||||
Db 841 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGACCAAGTACACCG 900

Qy 1058 CCTTCACCATCCCAGCATCAACAACGAGACCCCCGGCATCCGGTACCAGTACAACGTGC 1117
         |||||||||||||| |||||||||||||||||||
Db 901 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 960

Qy 1118 TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCAAGATCCTG 1177
         |||||||||||||||||||||||||||||||||
Db 961 TGCCCCAGGGCTGGAAGGGCAGCCCCAGCCATCTTCCAGAGCAGCAAGATCCTGG 1020

Qy 1178 AGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCCACCA------GGCCCCCCTGTACG 1231
         |||||||||||||||||||||||||||||||
Db 1021 AGCCCTTCCGCGCCCGCAACCCCGACATCGTGATCCACCACTACCAGTACACTGGACACGTACG 1080

RESULT 3
US-09-475-515-84
; Sequence 84, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 84
; LENGTH: 2312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(-).Protmod.RTopt(+)
; US-09-475-515-84

Query Match   81.3%; Score 1998.4; DB 4; Length 2312;
Best Local Similarity 92.9%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 141; Indels 24; Gaps 4;

```
Db   2161  ACAAGCTGGTGAGCGCCGGCATCCGCAAGGTGCTGTTCCTGAACGGCGGCGATCGATGGGCGGCA  2220
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   2366  TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCCAGCGGCAGCGGCGGCCCTAGGATCGATT  2425
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   2221  TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCCAGCGGCAGCGGCGGCCCTAGGATCGATT  2280

Qy   2426  AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC  2457
                  ||||||||||||||||||||||||||||||||

Db   2281  AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC  2312
```

RESULT 4
US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;   APPLICANT: BARNETT, Susan
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: SRIVASTAVA, Indresh
;   APPLICANT: LIAN, Ying
;   APPLICANT: HARTOG, Karin
;   APPLICANT: LIU, Hong
;   APPLICANT: GREER, Catherine
;   APPLICANT: SELBY, Mark
;   APPLICANT: WALKER, Christopher
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;   TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;   FILE REFERENCE: 1621.002
;   CURRENT APPLICATION NUMBER: US/09/475,515A
;   CURRENT FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 90
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 4319
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6

```
Query Match             78.0%;  Score 1915.8;  DB 4;  Length 4319;
Best Local Similarity   89.4%;  Pred. No. 6.8e-312;
Matches 2140;  Conservative   0;  Mismatches 227;  Indels  28;  Gaps   6;

Qy     14  TGGCCGAGGCCATGAGCCAGGCCACCAGC--GCCAACATCCTGATGCAGCGCAGCAACT   70
                 ||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db   1100  TGGCCGAGGCGATGAGCCAGGTGACGAACCCGGCCATCATGATGCAGCGGCGGCCAACT  1159

Qy     71  TCAAGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCC  130
                 |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1160  TCCGCAACCAGCCGAAGACCGAAGACCGTCAACTGCTTCAACTGCGGCAAGGAGGGCCACACCGCCA  1219

Qy    131  GCAACTGCGCGCCGCCCCCCCGCCAGAAGGGCTGCTGAAGCGGCTGGGCAAGGAGGGCCACCAGA  190
                 |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   1220  GGAACTGCCGCGCCGCCCCCCGCCGCAAGAAGGGCTGCTGGGCGGCCGCCGGAGGACACCAAA  1279

Qy    191  TGAAGGACTGCACCGAGCGCCAGGCCCAACTTCTTCCCGGGAGGGCCCTTCCCCCAGG  250
                 |||||||||||||||||||||||||||||||| ||||| ||||||
Db   1280  TGAAAGATTGCACTGAGCGACAGGCTAA-TTTTTTAGGGAAGATTCGGCCTTCCTTCCTACAAG  1338

Qy    251  GCAAGGGCCCGGAGTTCCCCAGCGAGCAGAACCGGCCAACAGCCCCACCAGCGCCGAGC  310
                 ||||||||||| ||||||||||||||||||||||||||||||||||
Db   1339  GGAAGGGCCAGGGAATTTTCTTCAGAGCAGCAGCAGCCAACAGCCCCACCAGAAGAGAGC  1398

Qy    311  TGCAGGTGCGGCGCG-----ACAACCCCCAGCGAGGCCGCGGCCGAGGGCCGCAGGGCA  364
                 ||||||||||||||       |||||||||||||||| |||||||||||||
Db   1399  TTCAGGTTTGGGGAGGAGAAAACAACTCCTCTCAGAAGCAGGCAGCAGATAGACAAGGAA  1458

Qy    365  ------CCCTGAACTTCCCCAGGATCACCCTGTGGCAGCGCATCAAGG  418
                       ||||||| ||||||||| |||||||||||||||||||
Db   1459  CTGTATCCTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCCTCGTCACAGTAAGGA  1518
```

```
Qy   1232  TGGGCAGCGCGACCTGGAGATCGGCCAGATCGGCCAAGATCGAGGAGCTGCGCAAGCACC  1291
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TGGGCAGCGACGGACCTGGAGATCGGCCAGATCGGCCAAGATCGAGGAGGAGCTGCGCAAGCACC  1140

Qy   1292  TGCTGCGGCTTCACCACCCCCGACAAGAAGACACCAGGAGACGCCCCCTTCCTGC  1351
                  |||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  TGCTGCGGGCTTCACCACCCCCGACACCCCCGACAAGAAGACACCAGGAGACGCCCCCTTCCTGT  1200

Qy   1352  CCAT------CGAGCTGCACCCCGACAAGTGACAAGTGGACCCATCGAGCTGCCCGAGA  1405
                  ||||       ||||||||||||| |||||||||||| ||||||||||
Db   1201  GGATGGGCTAGCGAGAGTGCGACCCTGCACCCGGCCCAGTGGGCCCCATCATGCTGCCCGAGA  1260

Qy   1406  AGGAGAGCTGACCCTGAACGAGCAATCCAGAGACATCCAGAAGCTGGGGCAAGCTGGGGCCAGCC  1465
                  ||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  AGGACAGCTGGACCCTGAACGAGCAATCCAGAGAAGCTGAACTGGGGGCAAGCTGGGGCCAGCC  1320

Qy   1466  AGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGGCCGCCAAGGCCC  1525
                  ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  AGATCTAGCGCCGGCATCAAGGTGAAGCACTGTGCAAGCTGCTGCGGCGCGGCCACCAAGGCCC  1380

Qy   1526  TGACCGACATCGTGCCCCTGACCGAGGAGGAGGCCGAGCTGGAGCTGGCCGAGACCGCGAGA  1585
                  ||||||||||||||| |||||||||||||||||||||||||
Db   1381  TGACCGAGGTGATCCCCCCTGACCGAGGTGGCCGAGGAGGGCCGAGCTGGGCCGAGAACCGCGAGA  1440

Qy   1586  TCCTTGCGCGAGCCGTGCACGGCGGCGTGTACTACGACCCCAGCAAGGAACCTGGTGGCCGAGA  1645
                  |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441  TCCTTGAAGGAGCCCGTGCACGAGGTGTACTACGACCCCAGCAAGGACCTGGTGGCCCGAGCC  1500

Qy   1646  TCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACC  1705
                  |||||||||||||||||||||||||||||||||||||||||||||
Db   1501  TCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACC  1560

Qy   1706  TGAAGACCGGCAAGTACGCCAAGATGCGACCCCGACCAACCAACGACTGAAGCAGTGA  1765
                  ||||||||| ||||||||||||||||||||||||||||||||||||
Db   1561  TGAAGACCGGCAAGTACGCCAAGATGCGCGGGCCCCCACCACCAACGACTGAAGCAGTGA  1620

Qy   1766  CCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCCAAGACCCCCAAGT  1825
                  ||||||||||||||||||||||||||||||||||||| |||||||||||
Db   1621  CCGAGGCCGTGCAGAAGGTGAGCACGGCATCGTGATCTGTGATCTGGGGCCAAGATCCCCAAGT  1680

Qy   1826  TCCCGCCTGCCATCCATCCAGAGAGGAGGAGACCTGGGACGACCTGGTGTTGCTACTGCGCAGCCA  1885
                  ||||||||||||||||||||||||||||||||||||||||||||||
Db   1681  TCAAGCTGCCCATCCATCCAGAGAGGAGGAGACCTGGGAGGCCTGGTGGAGTACTGCTGCGGTACCAGC  1740

Qy   1886  CCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCGGTGAAGCTGGTGGTACCAGC  1945
                  |||||||||||||||||||||||||||||||||||||||||||
Db   1741  TCAAGCTCCCCGAGTGGGAGTTCGTGAACACCCCCCCCGGTGAAGCTGGTGGTACCAGC  1800

Qy   1946  TGGAGAAAGAGCCCATCATGGGCGCCGAACCTTCTACGTGGACGGGCCGCCAACCCG  2005
                  ||||||||||||||||||||||||||||||||||||||||||||
Db   1801  TGGAGAAAGGAGCCCCATCATGTGGGGCGCCCAACCTTCTACGTGGACGGGCCGCCAACCGCG  1860

Qy   2006  AGACCAAGATCGGCAAGGCGGAGAGCGACGTCGACCGAGAAGATCGTGAGCC  2065
                  ||||||||||||||||||||||||||||||||||||||||||
Db   1861  AGACCAAGCTGGGCAAGGCTGGTGGTGACGCTGGTGGTGAGCA  1920

Qy   2066  TGACCGAGACCACACCAAGAAGCCAGAAGACCGAGCTGCAGGCCATCCAGCTGCCCCTGCAGGACACA  2125
                  |||||||||||||||||||||||||||||||||||||||||||||
Db   1921  TCGCCGACACCAGAGGTGACAACCAGCAGCAGGAAGGCTGCAGGGCCATCCAGCTGGCCCTGCCAGGACA  1980

Qy   2126  GCGCAGCAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC  2185
                  ||||||||||||||||||||||||||||||||||||||||
Db   1981  GCGGGCTGGAGGTGAACATCGTGGACGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC  2040

Qy   2186  AGCCCGACAAGAGCGAGAGCGAGGTCGTGAACCAGATCATCGAGCAGCTGATCAAGAAGG  2245
                  |||||||||| |||||||||||||||||||||||||||||||||||||||
Db   2041  AGCCCGACAAGAGCGAGAGCGAGGTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGG  2100

Qy   2246  AGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGGGGCATCGAGGCGAGATCG  2305
                  |||||||||||||||||||||||||||||||||||||||||||||
Db   2101  AGAAGGTGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGGCAGGGTG  2160

Qy   2306  ACAAGCTGGTGAGCGCCGGGCATCCGCAAGGTGCTGTTCCTGGACGGCGGCGATCGATGGCGGCA  2365
                  |||||||||||||||||||||||||||||||||||||||||||
Db   2306  ... 
```

Qy 1487 TGCGCCAGCTGTGCAAGCTGCTGCGGGCGGCCAAGGCCCTGACCGACGCATCGTGCCCCTGA 1546
        ||||||||||||||||||||||||||||||||| |||||||||||||||| ||||||
Db 2599 TGAAGCAGCTGTGCAAGCTGCTGCGGGCGGCCACCAAGGCCCTGACCGACCCCGTGA 2658

Qy 1547 CCGAGGAGGCCGGAGCTGGAGCTGGCCGAGATCCTGCGGCGAGAACCTGGAGGAGGCCGTGCACG 1606
        ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db 2659 CCGAGGAGGCCGGAGCTGGAGCTGGCCGAGATCCTGAAGCGAGGAGCCGTGCACG 2718

Qy 1607 GCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACC 1666
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2719 AGGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGGCC 2778

Qy 1667 AGTGGACCTACACCAGGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1726
        |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db 2779 AGTGGACCTACACCAGGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCC 2838

Qy 1727 AGATGCCACCGCCACCACCAACGACGTGAGGCAGCTGACCGAGGCCGTGCAGAAGATCG 1786
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2839 GCATGCCGCCGCCGCCGCCAACGACGTGAGGCAGCTGACCGAGGCCGTGCAGAAGGTGA 2898

Qy 1787 CCATGGAGAGCATCGTGATCTGGGACCGCAGGCCCACCTGGATCCCGAGTGGGAGT 1846
        ||||||||||||||||||| ||||||||||||| |||||||||||||| ||||||||||
Db 2899 GCATCGAGAGCATCGTGATCTGGGACCAAGGTCAAGCTGCCATCCCGAGTGGGAGT 2958

Qy 1847 AGACCTGGAGACCTGGTGGACCGACTACTGCCAGGCCACCTGGATCCCGAGTGGGAGT 1906
        |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db 2959 AGACCTGGAGACCTGGTGGATGGAGTACTGGATCCCGAGTGGGAGT 3018

Qy 1907 TCGTGAACACCCCCCCTGGTGAAGCTGTGGTGGTACCAGCTGGAGAAGGAGCCCATCATCG 1966
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3019 TCGTGAACACCCCCCCTGGTGAAGCTGTGGTGGTACCAGCTGGAGAAGGAGCCCATCGTGG 3078

Qy 1967 GCGCCGAGACCTTCTACGTGACGGCGCCGCCAACCGGAGACCAAGATCGGCAAGGCCG 2026
        |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db 3079 GCGCCGAGACCTTCTACGTGACGGCGCCGCCAACCGGAGACCAAGGTCGGCAAGGCCG 3138

Qy 2027 GCTACGTGACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGA 2086
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3139 GCTACGTGACCGACCGGGGCCGGCAGAAGATCGTGAGCATCGGCCGCACCACCAACCAGA 3198

Qy 2087 AGACCGAGCTGCAGGCCATCACGCAGCTCCAGCTGGCCTGGAGGTGAACATCG 2146
        ||||||||||||||||| |||||| |||||||||||||||||||||||||||
Db 3199 AGACCGAGCTGCAGGCCATCACCTGGGCCTGGAGGTGAACATCG 3258

Qy 2147 TGACCGACAGCCAGTACGCCCTGGGCATCATCAGGGCTGGAAGGGCA 2206
        |||||||||||||||||||| |||||||| ||||||||||||||||
Db 3259 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCGCGTGGAGGGCA 3318

Qy 2207 AGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAGGAGGAAGGTGTACCTGAGCTGG 2266
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3319 AGCTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAGGAGGAAGGTGTACCTGGCCTGG 3378

Qy 2267 TGCCCGCCCAAGGGCATCGATGGCCGCAACAAGGGCCGCGGCA 2326
        ||||||||||||||| ||||||||||| ||||||||||||||||
Db 3379 TGCCCGCCCAAGGGCATCGGCGGCAACAAGGGCCGCGGCA 3438

Qy 2327 TCCCGAAGGTGCTGTTCCTGGACGGCATCGATCGTGATCTACCAGTA 2381
        |||||||||||||||||||||| ||||||||||||||||||||||||
Db 3439 TCCCGAAGGTGCTGTTCCTGGACGGCATCGATCGTGATCTACCAGTA 3493

RESULT 5
US-09-475-515-81
; APPLICANT: BARNETT, Susan.
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine

Qy 419 TGGGCGGCCAGATCAAGGAGGAGGCCCTGCTGGACACCGTGCTGGAGG 478
       ||||||||||||||||||||||| |||||||||||||||||||||||
Db 1519 TCGGCGGCAGCTCAAGGAGGCGCTGCTGCTGCTGACACCGTGCTGGAGG 1578

Qy 479 AGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCGGCTTCATCA 538
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1579 AGATGAACCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGGGATCGGCGGGGCTTCATCA 1638

Qy 539 AGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCG 598
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1639 AGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCG 1698

Qy 599 TGCTGATCGGCCCCACCCCTGAACATCATCGGCGCGCGCAACATGCTGACCCAGCTGGGCT 658
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1699 TGCTGTGGTGGCCCCACCCCTGAACATCATCGGCGCGCGCAACCTGCTGACCCAGCTGGGCT 1758

Qy 659 GCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCA 718
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1759 GCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGGA 1818

Qy 719 TGGACGGCCCCCAAGGTCAAGCAGTGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG 778
       |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 1819 TGGAACGGCCCCCAAGGTCAAGCAGTGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGG 1878

Qy 779 CCATCTTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCCAAGATCGGCCCCCGAGAACCCCT 838
       |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db 1879 AGATCTGCACCGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCT 1938

Qy 839 ACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGG 898
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1939 ACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGG 1998

Qy 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 958
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1999 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 2058

Qy 959 ACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACT 1018
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2059 ACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACT 2118

Qy 1019 TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACACCGCCTTCACCATCCCCAGCATCA 1078
        |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db 2119 TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA 2178

Qy 1079 ACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA 1138
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2179 ACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA 2238

Qy 1139 GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAAACC 1198
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2239 GCCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACC 2298

Qy 1199 CCGAGATCGTGATCTACCA-------GGCCCCCGTGACGTGGCAGCGCACCTGGGAGATCG 1252
        ||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 2299 CCGACATCGTGATCTACCAGTACATGGACGACCTGTACGTGGCAGCGACCTGGGAGATCG 2358

Qy 1253 ACATCCAGAAGTGGTGGCAAGCTGGGCGCTGGAGCTGCTCGGCCCCGGCCATCAAGG 1312
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2359 ACATCCAGAAGTGGTGGCAAGCTGGGCGCTGGAGCTGCTCGGCCCCGGCCATCAAGG 2418

Qy 1313 CCCCCGACAAGAGCACCAGGAGGAGGGAGCCCC------CCGAGCTGCACC 1366
        |||||||||||||||||||||||||||||||| |||||||||||
Db 2419 CCCCCGACAAGAGCACCAGGAGGAGGGAGCCCCCCCTTCCTGTGGCAGGAGCTGGAGCTGCACC 2478

Qy 1367 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGACTGACCGGTGAACG 1426
        ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db 2479 CCGACAAGTGGACCGTGCAGCCCATCGTGCTGCCCGAGAAGGACAGCTGGACCGTGAACG 2538

```

```
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
;   LENGTH: 2299
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: FS(+).proinact.RTopt.YMWM
US-09-475-515-81

  Query Match   76.2%; Score 1872.8; DB 4; Length 2299;
  Best Local Similarity   89.3%; Pred. No. 9.6e-305;
  Matches 2055; Conservative   0; Mismatches 232; Indels   13; Gaps   3;

Qy   170 GCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCG 229
           ||||| || || ||||| ||||| || ||||| ||| | ||||| || || || | |
Db   1 GCGGCCGCGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTT-AGGG 59

Qy   230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCA 289
           | || |||||||||| || || |||||| | || || | ||||||| | | ||||
Db   60 AAGATCTGGCCTTCCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCA 119

Qy   290 ACAGCCCCACCAGCCGCGAGCTGCAGGTGCGCGGCG------ACAACCCCCGCAGCGAGG 343
           ||||||||||||| | ||||| ||||| | || | ||||| ||| | || |
Db   120 ACAGCCCCACCAGAAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACTCCCTCTCAGAAG 179

Qy   344 CCGGCGCCGAGCGCCAGGGCA------CCCTGAACTTCCCCCAGATCACCCTGTGGCAGC 397
           | || ||||| | || || | || | |||||||| |||||||| || ||||| |
Db   180 CAGGAGCCGATAGACAAGGAACTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAAC 239

Qy   398 GCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCG 457
           | ||||| || | || | | | || || || ||||||| || ||||| || || || |
Db   240 GACCCCTCGTCACAATAAGGATCGGGGGGCAACTCAAGGAAGCGCTGCTCGATACAGGAG 299

Qy   458 CCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCG 517
           | || || || || | || || |||| |||| || || ||||| || || ||||| |
Db   300 CAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGATAG 359

Qy   518 GCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCG 577
           | || ||||| ||||||||||||||| | |||||||||||||| | | || ||||| |
Db   360 GGGGGATCGGGGGCTTCATCAAGGTGAGGCAGTACGACCAGATACCTGTAGAAATCTGTG 419

Qy   578 GCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCA 637
           | | || || || || || | | | || || || || || ||||| || || | |
Db   420 GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAA 479

Qy   638 ACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGC 697
           | || |||||||| | ||||||||| ||||||||||||||||||| || ||||| ||||
Db   480 ATCTGTTGACCCAGATCGGCTGCACCTTGAACTTCCCCATCAGCCCTATTGAGACGGTGC 539

Qy   698 CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG 757
           |||||||| ||||||| || ||||||||||||||||| ||||| ||||| |||||||||
Db   540 CCGTGAAGTTGAAGCCGGGGATGGACGGCCCCAAGGTCAAGCAATGGCCATTGACCGAGG 599

Qy   758 AGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCA 817
           ||||||||||||||||| | |||||| ||||||||||||||||||||||||| ||
Db   600 AGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATCAGCA 659

Qy   818 AGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCA 877
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 AGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCA 719

Qy   878 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 937
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 779

Qy   938 AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC 997
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780 AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC 839

Qy   998 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCG 1057
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   840 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACAAGGACTTCCGCAAGTACACCG 899

Qy   1058 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 1117
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 959

Qy   1118 TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db   960 TGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1019

Qy   1178 AGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCA 1237
           ||||||||||| | |||||||| ||||||||||||||||||||||||||||||||||
Db   1020 AGCCCTTCCGCAAGCAGAACCCCGACATCGTGATCTACCAGGCCCCCCTGTACGTGGGCA 1079

Qy   1238 GCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1297
           |||||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db   1080 GCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGCCAGCACCTGCTGC 1139

Qy   1298 GCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGCCCATCG 1357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1140 GCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGCCCATCG 1199

Qy   1358 AGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGA 1417
           ||||||||||||||||||||||||||||||||||| ||||||||||||||| |||||||
Db   1200 AGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCATGCTGCCCGAGAAGGACAGCTGGA 1259

Qy   1418 CCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCG 1477
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1260 CCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACGCCG 1319

Qy   1478 GCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCG 1537
           ||||||||||| |||||||||||||||||||||||| |||||||||||||||| |
Db   1320 GCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGAGGTGA 1379

Qy   1538 TGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGC 1597
           | ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   1380 TCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGC 1439

Qy   1598 CCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGG 1657
           ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1440 CCGTGCACGAGGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGG 1499

Qy   1658 GCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCA 1717
           |||| | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1500 GCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCA 1559

Qy   1718 AGTACGCCAAGATGCGCACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGC 1777
           |||||||| |||||| |||||||||||||||||||||||||||||||||||||||||
Db   1560 AGTACGCCCGCATGCGCGGCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGC 1619

Qy   1778 AGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCA 1837
           ||||| | || ||||||||||||||||||||||||| |||||||||| |||||||
Db   1620 AGAAGGTGAGCACCGAGAGCATCGTGATCTGGGGCAAGATCCCCAAGTTCAAGCTGCCCA 1679

Qy   1838 TCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCG 1897
           |||||||||||||||||||| ||||||||| || |||||||||||||||||||||||||
Db   1680 TCCAGAAGGAGACCTGGGAGGCCTGGTGGATGGAGTACTGGCAGGCCACCTGGATCCCCG 1739

Qy   1898 AGTGGAGTTCGTGAACACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGC 1957
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1740 AGTGGAGTTCGTGAACACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGC 1799

Qy   1958 CCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCG 2017
           ||||| | |||||||||||||||||||||||||||||||||||||||||||| | |
Db   1800 CCATCGTGGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGCTGG 1859
```

```

```
Qy   60 AGATCTGCCTTCCTACAAGGGAAGGCCAGGAGGAATTTTCTTCAGAGCAGCAGCAGAGCCA 119
            ||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db 290 ACAGCCCACCAGCGCGAGCTGCAGGTGCCAGGTGCCGGCG------ACAACCCCCAGCGAGG 343

Qy 120 ACAGCCCCACCAGGAAGAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACTCCCTCTCAGAAG 179
            ||||||||||||||||||||||||||||
Db 344 CCGCGCGCAGCGCCAGGGCA------CCCTGAACTTCCCCCAGTACCCTGTGGCGAGC 397

Qy 180 CAGGAGGCCGATAGACAAGGAACTGTATCGTTTAACTTCCTCAGATCACTCTTTGGCAAC 239

Db 398 GCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGTCAAGGAGGGCCCTGGACACCGGCG 457

Qy 240 GACCCCTCGTCACATAACGATCGGGGGGGAGATGAATGCTTCGCCGGCCAAGATGATCG 299

Db 458 CCGGACGCACCGTGCTGAGGCAGCATCGGCGGCAACGACAGCTGGACAGCTGGTA 517

Qy 300 CAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAAATGGAAACCAAAAATGATAG 359

Db 518 GCGGCATCGCGGGCTTCATCAAGGTGCGCAGTACGACCCAGATCCTGATCGAGATCGTGCG 577

Qy 360 GGGGGATCGGGGGCTTCATCAAGGTGAGCGGCTTCATCAAGGTGCGCGGCATCGTGATCTACC 419

Db 578 GCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGAACATCATCGGCGCCA 637

Qy 420 GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCACATAATTGGAAGAA 479

Db 638 ACATGCTGACCCAGCTGGCTGCACCCTGAACTTCCCATCAGCCCCATTGAGACCGTGC 697

Qy 480 ATCTCGTTGACCGATCGATCGCTGCACCTTCGAACTTCCCATCAGCCCCATTGAGACGTGC 539

Db 698 CCGTGAAGCTGAAGGCCCGGGCCCGGCCCGGCCTGAAGCAGTGAAGCAGTGGCCCCTGACCAGG 757

Qy 540 CCGTGAAGTTGAACGCCCGGGGGGATGGAACCGGGACGGGCCCAAGGTCAAGCAAGTCAAGTTGACCGAGG 599

Qy 758 AGAAGATCAAGGCCCTGACCGGCCATCTCGCCGAGGAGATGGAGAGGAGGAGGAGGCAAGATCACCA 817
            |||||||||||||||||||||||||||
Db 600 AGAAGATCAAGGCCCTGGAGATCTGACCGAGATCTGACCCAGATGGAGAGGAGGAGGAGGCAAGATCAGCA 659

Qy 818 AGATCGGCCCCGAGAACCCTACAACACCCCGTGTTGCCATCGAAGAGAAGGACAGCA 877

Db 660 AGATCGGCCCCGAGAACCCTACAACACCCCCGTGTTGCCATCGAAGAGAAGGACAGCA 719

Qy 878 CCAAGTGGCGCAAGTGGTGAGCTTCCCGCAGTTCCGCAGTTGAACAAGCGCACCCAGGACTTCTGGG 937

Db 720 CCAAGTGGCGCAAGTGGTGAGCTTCCCGCAGTTCCGCAGTTGAACAAGCGCACCCAGGACTTCTGGG 779

Qy 938 AGGTGCAGCTGGGCATCCCCACCCCGCGCGCCGCTGAAGGAAGAAGAGCGTGACCGTGC 997

Db 780 AGGTGCAGCTGGGCATCCCCACCCCGCCCACCCCGGCCTGAAGGAAGAAGAGCGTGACCGTGC 839

Qy 998 TGGACGTGGGCGACGGCCTACTTCAGCGTGCCCCTGGACAAGTACACCG 1057

Db 840 TGGACGTGGGCGACGGCCTACTTCAGCGTGCCCCTGGACAAGTACACCG 899

Qy 1058 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGATCCAAGATCCTGG 1117

Db 900 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGATCCAAGATCGTGC 959

Qy 1118 TGCCCCAGGGCTGAAGGCTGAAGGCCAGCCCGCCATCTTCCAGAGCAGCATGCAGCCAAGATCCTGG 1177

Db 960 TGCCCCAGGGCTGAAGGCTGAAGGCCAGGGCCAGCCCGCCATCTTCCAGAGCAGCAGCATGCAGCCAAGATCCTG 1019

Qy 1178 AGCCCTTCCGGCCGCCAAGCACCCCCGAGATCGTGATCTACAGCCCCCCGTGTACGTGGGCA 1237

Db 1020 AGCCCTTCCGGCCGCCAAGCAGCACCCCGACATCGTGATCTGATCTGATCAAGATCGGTGTACGTGGGCA 1079

Qy 1238 GCGACCTGGAGATCGGCCCAGCACCGGCCACCGGCCGCCAAGACCCCAACTTCTTCCCGCG 1297

Db 1080 GCGACCTGGAGATCGGCCCAGCATCGGCCACCAAGACCGCCGAGCTGGCCAAGCACCTGCTGC 1139

Qy 1298 GCTGGGGCTTCACCCACCCCGCCGGCCAAGGCCGCGAGTTCCCCCAGTTCGCGAGAACCGCCGCA 1355

Db 1140 GCTGGGGCTTCACCCACCCCGCCGGCCAAGAAGCACCCCGAGAACCCCCGCCCCCTTCCTGTGGATGG 1199
```

```
Qy 2018 GCAAGGCCGGCTGGTACGTGACCGACCCGGGGCCGCGCGGGCGCCAGAAGATCGTGAGCGCCTGAGCGAGACCA 2077
Db 1860 GCAAGGCCGGCTGGTACGTGACCGACCCGGGGCCGCGCGGGCGCCAGAAGATCGTGAGCGCCTGACCGACCACCA 1919

Qy 2078 CCAACCAGAAGAGACCGGAGCTGCAGGCCATCGGGGCCATCGGCCTGGCCCCTGCAGCAGCAGCAGCGGCGCGAGG 2137
Db 1920 CCAACCAGAAGAGACCGGAGCTGCAGGCCATCGGCCTGGCCCCTGCAGCAGCAGCAGCGGCGCTGGAGG 1979

Qy 2138 TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCGGGCATCATCCAGGCCCGACAAGA 2197
Db 1980 TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCATCATCCAGGCCCGACAAGA 2039

Qy 2198 GCGAGAGCGAGCTGGTGAACCAGAAACATCGAGCAGCTGATCAAGAGAGAAGGAGAAGGTGTACC 2257
Db 2040 GCGAGAGCGAGCTGGTGAAGCATCGAGCAGCTGATCATCGAGAGAAGGAGAAGGTGTACC 2099

Qy 2258 TGAGCTGGTGGTCGCCCGCCCACAAGGGCATCGGCGGCAACGACCAGTGGCAGAAGCTGGTGA 2317
Db 2100 TGGCCCTGGGTCGCCCGCCCACAAGGGCATCGGCGGCAACGACCAGTGGCAGAAGCTGGTGA 2159

Qy 2318 GCAAGGGCCATCCGCAAGGTGCGTGTTCCTGGACGTGGTCTTCCTGGACGTGATCTGATCTACC 2377
Db 2160 GCGCCGGCATCCGCAAGGTGCGTGTTCCTGGACGTGGTCTTCCTGGACGTGATCTGATCTACC 2219

Qy 2378 AGTACATGACGACCACTTGTACGTGGGCAGCGGGCGGCGGCGGCCTAGGAATCGATTGAATTAAAAGCTTCCCG 2437
Db 2220 AGTACATGACGACCACTTGTACGTGGGCAGCGGGCGGCGGCGGGGGCCTAGGAATCGATTAAAAGCTTCCCG 2279

Qy 2438 GGGCTAGCCACCGGTGAATTC 2457
Db 2280 GGGCTAGCCACCGGTGAATTC 2299
```

RESULT 6
US-09-475-515-80
; Sequence 80, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;_ APPLICANT: BARNETT, Susan
;_ APPLICANT: ZUR MEGEDE, Jan
;_ APPLICANT: SRIVASTAVA, Indresh
;_ APPLICANT: LIAN, Ying
;_ APPLICANT: HARTOG, Karin
;_ APPLICANT: LIU, Hong
;_ APPLICANT: GREER, Catherine
;_ APPLICANT: SELBY, Mark
;_ APPLICANT: WALKER, Christopher
;_ TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;_ TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;_ FILE REFERENCE: 1621.002
;_ CURRENT APPLICATION NUMBER: US/09/475,515A
;_ CURRENT FILING DATE: 1999-12-30
;_ NUMBER OF SEQ ID NOS: 90
;_ SOFTWARE: PatentIn Ver. 2.0
;_ SEQ ID NO 80
;_ LENGTH: 2305
;_ TYPE: DNA
;_ ORGANISM: Artificial Sequence
;_ FEATURE:
;_ OTHER INFORMATION: Description of Artificial Sequence:
;_ OTHER INFORMATION: FS(+).proinact.RTopt.YM
US-09-475-515-80

Query Match   75.4%; Score 1852; DB 4; Length 2305;
Best Local Similarity 89.0%; Pred. No. 2.9e-301;
Matches 2052; Conservative 0; Mismatches 235; Indels 19; Gaps 4;

```
Qy  170 GCGCCAAGGAGGAGGCCACCGGGCCAGATGAAGGACTGCACCGAGCGCCAGGCGCCAACTTCTTCCGCG 229
            ||  || |||||||||        |  ||||           |||||  |||||||||||||
Db    1 GCGCCCGCGAAGGAGACACCCCAAATGAAGATTGCACTGAGAGATTGCACTGAGAGACAGGCTAATTTTTT-AGGG 59

Qy  230 AGGAGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGGAGTTCCCCCAGCGAGCGAGACCGCCGCCA 289
            ||   ||| || |||  |||||     ||||||||||  || ||
Db
```

Qy　2432　TTCCCGGGGCTAGCACCGGTGAATTC　2457
Db　2280　TTCCCGGGGCTAGCACCGGTGAATTC　2305

RESULT 7
US-09-552-950-2
; Sequence 2, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon
; OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

Query Match　　　　66.2%;　Score 1626.6;　DB 4;　Length 4307;
Best Local Similarity　82.2%;　Pred. No. 1.4e-263;
Matches 1951;　Conservative　0;　Mismatches 394;　Indels　28;　Gaps　6;

Qy　　12　CATGGCGAGCCATGAGCCAGG---CCACCAGCGCCAACATCCTGATCGCAGCGCAGCAA　68
　　　　　　||||||||||||||||||||　　||||||||||||||||||||||||||||||||||||
Db　1086　CCTGCTGAGGCCATGAGCCAGGTGACCAGTGCAGGTGACCAACTCCGTCCATCATGATGCAGCGCAGCGGCAA　1145

Qy　　69　CTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGC　128
　　　　　　||||||||||||||||||||||||||||||||||||||||||
Db　1146　CTTTCGGAACCAACCAAGATCGTCAAGTGCTTCAACTGTGGCAAAGAAGGGCACACAGC　1205

Qy　129　CCGCAACTGCGCGGCCCGCCCCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCA　188
　　　　　　||||||||||||||||||||||||||||||||||||||||||||||
Db　1206　CCGCAACTGCAGGGAGCCCCTAGGAAAAAAGGGCTGCTGGAAATGCGGCAAGGAAGGCCACCA　1265

Qy　189　GATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCA　248
　　　　　　||||||||||||||||||||||||||||||||
Db　1266　GATGAAAGACTGTACGGAGCAGCAGGCTAA-TTTTTTAGGGGAAGATCTGGCCTTCCTACA　1324

Qy　249　GGGCAAGGCCCGCGAGTTCCCCGCAGAGAACCGCGCCAACAGCCCCACCAGCCCGGCGA　308
　　　　　　|||||||||||||||||||||||||||||||||||||||
Db　1325　AGGGAAGGCCAGGGAATTTTCTTCAGCAGCAGCAGCAGCCCCAACAGCCCCACCGCCGCGA　1384

Qy　-309　GCTGCAGGT------GGGCGGCGACAACCCCCGGCGGCGAGGCCGGCGCCCGAGCGCAGGG　362
　　　　　　||||||||　　　　||||||||||||||||||||||||||||||||||||||
Db　1385　GCTTCAGGTCTGGGGTCGCGACAACAACTCCCCCTCCGAGCAGGAGCCGCGGCAGGG　1444

Qy　363　CA------CCCTGAACTTCCCCAGATCCCCTCAGTGACCGTTGGACCATCAA　416
　　　　　　||　　　　　||||||||||||||||||||||||||||||||||||
Db　1445　CACGGTGTCCTTCCTTCAACTTCCCTCAGTCCATCCAGCCCTTGGCAGGCACCGTCACCATCAA　1504

Qy　417　GGTGGCGGCGGGCAGGGCCCTGCTGACACCGTGCTGGA　476
　　　　　　||||||||||||||||||||||||||||||||||||||
Db　1505　GATCGGGGCGGGCAGCTCAAGGAGGGCTTCTCGAGCAGACGCAGCACCGTGCTGGA　1564

Qy　477　GGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGCGGCGGCTTCAT　536
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||
Db　1565　GGAGATGGTCGTTGCCAGGCGGCTGGAAGCCGAAGATGATCGGCGGAATCGGCGGTTTCAT　1624

Qy　537　CAAGGTGCGCGCCAGTACGACCAGATCCTGATCGGAAGGCCATCGGCAC　596
　　　　　　||||||||||||||||||||||||||||||||||||||||||||
Db　1625　CAAGGTGCGCGCCAGTATGACCAGATCCTCATCGGAAATCTGCGCCCGCCCAAAGGCTATCCGTAC　1684

Qy　597　CGTGCTGATCGGCCTCCCCACCACCCCCCACCCGTCAACATCGCTGACCCAGTCGGG　656
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||||
Db　1685　CGTGCTGGTGGGCCCCACCACCCCCCTGACCAACATCGCTGTTGACGCAACCTGTTGACGCAGATCGG　1744

Qy　1356　----CGAGCTGCACCCGACAAGTGACGTGCAGCCCATCGAGCGTGCCCGAGAGGAGA　1411
　　　　　　　　　||||||||||||||||||||||||||||||||||||||||||
Db　1200　GCTACCGAGCTGCACCCCGACAAGTGACCGGTGACCCGTGCAGCCCATCATGTGCCCGAGAGGACA　1259

Qy　1412　GCTGGACCGTGAACGACATCCAGAAGCTGGGCAAGCTGAACTGGGCCGCCAGATCT　1471
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||||
Db　1260　GCTGGACCGTGAACGACATCCAGAAGCTGGGCAAGCTGAACTGGGCCGCCAGATCT　1319

Qy　1472　ACCCCGCATCAAGGTGCGCCAGCTGCTGCGCGGCGGCCAAGGCCCTGACCG　1531
　　　　　　||||||||||||||||||||||||||||||||||||||||||
Db　1320　ACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCGCACCCAGGCCCTGACCG　1379

Qy　1532　ACATCGTGCCCCTGACCGAGGAGGCCGGAGACCTGGCCGAGAACCGGCGAGATCCTGC　1591
　　　　　　||||||||||||||||||||||||||||||||||||||||||||||||
Db　1380　AGGTGATCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGA　1439

Qy　1592　GCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGA　1651
　　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1440　AGGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGA　1499

Qy　1652　AGCAGGGCCACGACCAGTGACCTGGACCTACCAGGACGCCCTTCAAGAACCTGAAGA　1711
　　　　　　||||||||||||　　||||||||||||||||||||||||||||||||||
Db　1500　AGCAGGGCCAGGGCCAGTGACGACCTACCACAGGACGCCCTTCAACGAGCCCTTCAAGAACCTGAAGA　1559

Qy　1712　CCGGCCAAGTACGCCCAAGAGTGCGCACCCCCACCAACGACCTGAAGCAGCTGACCGAGG　1771
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||
Db　1560　CCGGCCAAGTACGCCCGCATGCGCGGCGCCCACACCAACGACGTGAAGCAGCTGACCGAGG　1619

Qy　1772　CCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCACCCCCAAGTTCCGCC　1831
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||
Db　1620　CCGTGCAGAAGTGAGCACCGAGGACATCGTGATCTGGGGCAAGATCCCCAAGTTCAAGC　1679

Qy　1832　TGCCCATCCAGAAGGAGACCTGGGACACCTGGTGGGACCGACTACTGCAGGCCACCTGGA　1891
　　　　　　||||||||||||||||||||||||||||||||||||||||||||||||
Db　1680　TGCCCATCCAGAAGGAGACCTGGGATGGAGGCTGGTGGATGAGTACTGGCAGGCCACCTGGA　1739

Qy　1892　TCCCGAGTGGGAGTTCGTGAACACCCCCCCCCCTGTGTAGCTGGTGTACCAGCTGTGAGA　1951
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||
Db　1740　TCCCGAGTGGGAGTTCGTGAACACCCCCCCCCCCCTGGTGAAGCTGTGTGTACCAGCTGTGAGA　1799

Qy　1952　AGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCA　2011
　　　　　　||||||||||||||||||||||||||||||||||||||||||||||||
Db　1800　AGGAGCCCATCGTCGGGGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCA　1859

Qy　2012　AGATCGGCAAGGCCGGCTACGTGACGAACGCCGGCGGGGGCCGGCAGAAGATCGTGAGCCTGACCG　2071
　　　　　　|||||||||||||||||　|||||||||||||||||||||||||||||||
Db　1860　AGTCTGGCAAGGCCGGCTACGTGACCAACGCCGGCCGCCAACAGGTCGTGAGCATCGCCG　1919

Qy　2072　AGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGCCCTGCAGGACAGCGGGCA　2131
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||
Db　1920　ACACCACCAACCAGAAGACCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACAGCGGCC　1979

Qy　2132　GCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCCTGGGCATCATCGAGGAAGAGGCCCCCAGCCG　2191
　　　　　　|||||||||||||||||||||||||||||||||||||||||||
Db　1980　TGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCGAGGCCCAGCCCGG　2039

Qy　2192　ACAAGAGGAGCGAGGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAGG　2251
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||||
Db　2040　ACAAGAGCGAGCAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG　2099

Qy　2252　TGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGC　2311
　　　　　　||||||||||||||||||||||||||||||||||||||||||||||||
Db　2100　TGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGC　2159

Qy　2312　TGGTGAGCAAGGGGCGGCCGCAAGGTGCGCGAAGGTCGTGTTCCTGGACGGCATCGATGGCGGCATCGTGA　2371
　　　　　　||||||||||||||||||||||||||||||||||||||||||||||
Db　2160　TGGTGAGCGCCGGCATCCGGAAGGTGCTGTTCCTGGACGGCATCGATGGCGCCATCGTGA　2219

Qy　2372　TCTACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGATTAAAAGC　2431
　　　　　　|||||||||||||||||||||||||||||||||||||||||||||
Db　2220　TCTACCAGTACATGGACGACCTGTACGTGGGCTCCGACCTGGAGATCGGATTAAAAGC　2279

RESULT 8
US-09-552-950-5
; Sequence 5, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
; LENGTH: 9772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:pSYNGP
US-09-552-950-5

Query Match   65.1%;   Score 1599.4; DB 4; Length 9772;
Best Local Similarity   81.5%;   Pred. No. 5.5e-259;
Matches 1934; Conservative   0; Mismatches 411; Indels   28; Gaps   6;

Qy   12 CATGGCGCAGGCCATGAGCCAGG---CCACCAGGCGCCAACATCCTGATGCAGCCGCAGCAA   68
               ||||||||||||||||||||||| || ||||||||||||| ||||||| |||||||||||
Db   2193 CCTGGCTGAGGCCATGAGCCAGGTGACCAACTCCGTACCATCATCATGATGATGCGCGGCAA 2252

```
Db    2276  TCGGGGGCAACTCAAGAGGAGCGGCTGCTCGATACAGGAGCAGCAGATGATACAGTATTATTAGAAG  2335

Qy     479  AGATGAGCCTGCCCGCGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGGCTTCATCA  538
                                                                                |||||||
Db    2336  AAATGAGTTTGCCAGAAGATGGAAACCAAAAAATGATAGGGGGGATCGGGGGGCTTCATCA  2395

Qy     539  AGGTGCGGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGCGCACCG  598
                                                                                ||||||
Db    2396  AGGTGAGCAGTACGACCAGATACTCATAGAAAATCTGTGACATAAAGCTATAAGTACAG  2455

Qy     599  TGCTGATCGGCCCCACCCCCGTGAACATCATCGGCCAACATGCTGACCCAGCTGGGCT  658
                                                                                |
Db    2456  TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATCGGCT  2515

Qy     659  GCACCCTGAACTTCCCATCAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCA  718
                                                                                ||||||||
Db    2516  GCACCCTGAACTTCCCCATCAGCCCTTATTGGACGCGGTGCCCGTGAAGCCCGGGA  2575

Qy     719  TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGGAAGAGATCAAGGCCCTGACCG  778
                                                                                |||||||
Db    2576  TGGACGGCCCCAAGGTCAAGCAATGGCCATTGACGAAAGAGAGATCAAGGCCTTAGTCG  2635

Qy     779  CCATCTGCGGAGGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCT  838
                                                                                |
Db    2636  AAATCTGTACAGAGATGGAGAGGAGGAGGAGTCAGCAAGATCGGGCCTGAGAACCCCT  2695

Qy     839  ACAACACCCCGTGTTCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG  898
                                                                                |
Db    2696  ACAACACTCCAGTCTTCCCAATCAAGAAGAAGGACAGTACCAAGTGGAAAGCTGGTGG  2755

Qy     899  ACTTCCGCGAGCTGAACAAGCCGCACCCAGGACTTCTCGGAGCGTGCAGCTGGCATCCCC  958
                                                                                |||||||
Db    2756  ACTTCAGAGAGCTGAACAAGCCGAACTCAGGACTTCTCGGGAAGTTCAGCTGGCATCCCAC  2815

Qy     959  ACCCCGCCGCGGCCTGAACGAGCGAAGAGGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCTACT  1018
                                                                                ||||||
Db    2816  ATCCCGCTGGGTTGAAGAGAAGTCAGTGACAGTGCATGCTCTGGATGTGGTGATGCCTACT  2875

Qy    1019  TCAGCGTGCCCTGGACGAGGACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCA  1078
                                                                                |||
Db    2876  TCTCCGTTCCCTTGGACGAGGACTTCAGGAGTTCAGGAAGTACACTGCCTTCACGATACCTAGCATCA  2935

Qy    1079  ACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCAGGGCTGGAAGGGCA  1138
                                                                                |||||||
Db    2936  ACAACGAGACACCAGGCATCCGCTACCAGTACAACGTCTGCCAGCAGGGATGGAAGGGAT  2995

Qy    1139  GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCCAACC  1198
                                                                                ||||||
Db    2996  CACCAGCCATCTTTCAAAGCAGCATGACCAAGATCCTGGAGCCTTCCGCCAAGCCGCCA  3055

Qy    1199  CCGGAGATCGTGATCTACCA-------GGCCCCCCTGTACGTGGCAGCCTGCGAGATCG  1252
                            ||||||||
Db    3056  CAGACATCGTGATCTATCAGTACATGGAGCTGCGAGAAGTGTAGGACTGACCTGAGATCG  3115

Qy    1253  GCCAGCACCCGCCAAGATGCTGCGCCAAGCACCTGCTTGCGCGTTCACCA  1312
                                                                                |||
Db    3116  GGCAGCACAGACAGGGCCAAGATACCAGATCGGAGGACGTGAGACTGTTGAGGTGGGGACTGACCA  3175

Qy    1313  CCCCCGACAAGAAGCACCAGAAGGAAGCACCCCCTTCCTGCCCAT-----CGAGCTGCACC  1366
                                                                                |||
Db    3176  CACCAGACAAGAAGCACCAGAAGGAACCTCCCTTCCTCCTGTGGATGGGCTACGAACTGCATC  3235

Qy    1367  CCGACAAGTGACCGTGCAGCCCATCGAGCTGCCCAGGAAGAGCAGCTGGACGTGGACGA  1426
                                                                                |||||||
Db    3236  CTGACAAGTGACAGTGCAGCGCCCCATCGGCTGTGCTGTCGCCCTGAGAAGAGCAGCTGGACTGTGAACG  3295

Qy    1427  ACATCCAGAGCTGTGGGGCAAGTGAACTGGGCGCAGCCCGGCATCAAGG  1486
                                                                                ||||||
Db    3296  ACATACAGAGCTCGTGGGCAAGTGAACTGGGCAAGCCAGGCATCAAAG  3355

Qy    1487  TGCGCCAGTGTGCAAGGACGCGTGCTGCGCGAGCGACCGTCTGCCCCTGA  1546
                                                                                |||
Db    3356  TTAGGCAGCTGTGCAAGTGCTTCGAGGAACAGGCACTGACAGAAGTGATCCCACTGA  3415
```

RESULT 9
US-09-872-733A-6
```
Db    4412  CGAGCTGGTCAATCAGATCAGATCATCGAGCAGCTGATCAAGAGGAGGAAAGGTTCTATCTGGCCTG  4471
                                                                                |
Qy    2265  GGTGCCCGCCCACAAGAGGCATCGGCCGCGGCATCGGCCGCGCAACGAGCAGCTGTGGTGAGCAAGGG  2324

Qy    2472  GGTACCCGCCCACAAAGGCATTGGCGCGGCAATGAGCAGGTCGACAAGCTGGTCTCGGCTGG  4531

Qy    2325  CATCCGCAAGGTGCTGTTCCTGGACGCGCATCGA  2357
                     |||||||
Db    4532  CATCAGGAGAGGTGTGCTATTCCTGGATGCGCATCGA  4564
```

; Sequence 6, Application US/09872733A
; Patent No. 656706
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO: 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBNKan containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733A-6

```
Query Match             62.6%;  Score 1539.2;  DB 4;  Length 8366;
Best Local Similarity   79.9%;  Pred. No. 6.2e-249;
Matches 1896;  Conservative   0;  Mismatches  448;  Indels  28;  Gaps   6;

Qy      14  TGGCCGAGGCCATGAGCGCAGGCCACCAGC--GCCAACATCCTGATCGAGCGCAGCAACT  70
                                                                                |||||
Db    1857  TGGCCGAGGCCGATGAGCCAGGTGACGAACTCGGCGACCATTAATGATGCAGAGAGGCAACT  1916

Qy      71  TCAAGGCCCAAGCGCAGATCGTTCAAGTGCTTCAAACTGCTTCAATTGTGCAAGACATCGCCC  130
                                                                                ||||
Db    1917  TCCGGAACCAGCGCAGCCGCGCCCGGGGAAGAAGGGCAAAGAGAGGCCACCGCCA  1976

Qy     131  GCAACTGCCGCGCGGCCGCCCCGGGAAGAAGGGCCAAGGTGGGCCAAGGAGGGCCACCAGA  190
                                                                                |||
Db    1977  GGAACTGCCGCGGGCTGTGTGAAATGTGGAAAATGTGGAGAAGGAAGGAACACCAAA  2036

Qy     191  TGAAGGACTGACGAGGCCCGCCGAGTTCCCGCGCAGAGAACCGCGCCAGCGGACCGCCAACC  250
                                                                                |||
Db    2037  TGAAAGATTGTACTGAGAGACAGGCTAA-TTTTTAGGGAAGATCTGGCCTTCCTACAAG  2095

Qy     251  GCAAGGGCCGCCGGCAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCAGCCGCCGCAGC  310
                                                                                ||||||
Db    2096  GAAGGGCCAGGAGAATTTCTTCAGAGCAGCAACAACTCCCCTCCAGAGCCCACCAGAAGAGAGC  2155

Qy     311  TGCAGGTGCGCGG------CGACAACCCCGCGAGGGCCGGCCCGAGCCCAGGGCA  364
                                                                                |
Db    2156  TTCAGGTTCTGGGGGTAGAGACAACACCCGTGGCAGCCGCCCCTGGTGAGCAGCATCAAGG  2215

Qy     365  ------CCCTGAACTTCCCCAGATCACCCCGATCACTCTTTGGCAACGACCCCCTCGTCACAGTAAGGA  418
                                                                                |||
Db    2216  CTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCCTCGTCACAGTAAGGA  2275

Qy     419  TGGCGGCCAGATCAAGGAGCCCTGCTCAAGGACACCGCCGACGACCCGTGCTGCTGAGG  478
                     |||||||
Db    3356  TTAGGCAGCTGTGCAAGTGCTTCGAGGAACAGGCACTGACAGAAGTGATCCCACTGA  3415
```

```
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 1
;  LENGTH: 4338
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
;  OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733A-1

Query Match             61.2%;  Score 1503.8;  DB 4;  Length 4338;
Best Local Similarity   79.8%;  Pred. No. 4.7e-243;
Matches 1896;  Conservative   0;  Mismatches 447;  Indels 34;  Gaps  9;

Qy    14 TGGCCGAGGCCATGAGCCAGGCCACCAGC---GCCACATCCTGATGCAGCCAGCAACT 70
         |||||||||| |||||||||||  || ||    || ||  |||| |||||| || ||||
Db  1085 TGGCCGAGGCGATGAGCCAGGTGACGAACTCGGCGACCATAATGATGCAGAGGGCAACT 1144

Qy    71 TCAAGGGCCCCAAGCGCCATCATCAAGTTCGCGGCAAGCTCAACTCGCGGCAAGAGGGC 130
         || | |||||||||| || || ||||| | ||  |||| | |||| || ||| || ||| 
Db  1145 TCCGGAACCAGCGGAAGATCGTCAAGTGCTTCAAGTGCGGCAAAGAGAGGGCACACCGGCCA 1204

Qy   131 GCAACTGCCGCGCCCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGTGGGCCACCAGA 190
         | |||||||| ||||| ||| ||||||||||| ||||| | || ||   |        || ||| 
Db  1205 GGAACTGCCGGCGCCCCCGGAAGAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAA 1264

Qy   191 TGAAGGACTGCACGGAGCGCCAAGGCCAACTTCTTCCCGGAGAACCTGCCTTCCCCAGG 250
         |||| || ||| | |  |  || |  || | ||    |  || | ||| ||||| || |  
Db  1265 TGAAAGATTGTACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTCGCCTTCCTACAAG 1323

Qy   251 GCAAGGCCCGAGTTCCCGGAGTGAACCGCGGCCAACAGCCCGCAGCGCCCCCACCCAGAGAGC 310
         | ||| | | ||||  | ||   |||| | || || ||| | ||| |||  || | |||||||| 
Db  1324 GGAGGCAGGGGAATTTTCTTCAGACGAGAGCCAACAGCCCCACCAGCCCCACCAGAAGAGAGC 1383

Qy   311 TGCAGGTGCGCCG------CGACAACCCCCGCGAGCGGGAGCGCCGAGCGCGCCAGGGCA 364
         |   |||||| |        |||| |||| |  ||| |  ||| ||||| ||| |||| |
Db  1384 TTCAGGTCTGGGGTAGAGACAACAACAACTCCCCCTCAGATCACCCTCGTGGCAGCAAGGAA 1443

Qy   365 ------CCTGAACTTCCCTCCCTCAGATCACCCTCCTTTTGGCAGCCATCTAAGG 418
               |||||||| | ||||  ||| | |   | ||   | |||  || || ||
Db  1444 CTGTATCCTTTAACTTCCCTCAGATCACTACTCTTTTGGCAGCGACCCGTCAACGTAAGGA 1503

Qy   419 TGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGG 478
         ||||||||||| |||||  || | || | || || | ||| | ||| || ||| || ||| 
Db  1504 TCGGGGGCAACTCAAGGAGCCGCTGTCGATCACAGGAGCAGATGATACAGTATTAGAAG 1563

Qy   479 AGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCA 538
         || | ||||| |||  ||  || ||||| ||| | ||| ||  ||| | |||  ||||  
Db  1564 AAATGAGTTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCA 1623

Qy   539 AGGTTCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGGCACAAGGCCATCGGCACCG 598
         || | | || |||| ||| |||  |    ||| |  |||  || |||| ||| ||| ||  
Db  1624 AGGTGCGACAGTACGAGATCCTGATCGGAGATGGGGACATAAAGCTATAGTACAG 1683

Qy   599 TGCTGATCGGCCCACC----CCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTG 654
         || | ||||||| |||     || | ||||||| ||   || | || | |||| ||||  
Db  1684 TATTAGTAGGACCTACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATC 1743

Qy   655 GGCTGCACCGTGACCCTGAACTTCCCCATCAAGCCCATCATGACCATCATCGGCCGCCCTG 714
         |||| ||||| |||| |||| |||||| ||||| || || ||| || || | | || |||  
Db  1744 GGCTGCACCTTGAACTTCCCCATCAGCCCTATTGAGACAGGTGCCGGTGCCGGTGAAGCCG 1803

Qy   715 GGCATGGACCGGCCCCAAGTGAAGTGGCAGTGCCCCGAGCAGTGGAAGATCAAGGCCCTG 774
         |||| ||||| || |||| || |||  || ||  | |||| |  || || || |||| |||  
Db  1804 GGGATGGACGGGCCCCAAGGTCAAGTCATTGGCCATTATGGCAGTAAATGAAGAGATCAAGGCCCTA 1863

Qy   775 ACCGGCCATCTGCGACGGATGGAGGAGGCCAAGATCACCAGGAGGCCCGAGAAC 834
         ||||| || || |  |||||| ||| | || || || |||  ||||||| ||||  
Db  1864 GTCGAAATCTGTACAGATGGAAGGAAGGAAGGAAGATCAGCAAGATCGGGGCTGAGAAC 1923

Qy   835 CCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTG 894
         ||||||| | |||| || | |||||| |||||| |   || ||| || ||| ||| |||
```

```
Qy  1547 CCGAGGAGGCCGAGCTGGCCGAGAACCCGCGCGGAGCCGTGCACG 1606
         | |||| || || ||| |  ||||||| | | ||| ||| |  ||
Db  3416 CAGAGGAAGCAGAGCTAGAACTGGCAGAACCGAGAGATCCTGAAGGAGCAGCCAGTACATG 3475

Qy  1607 GCGGTGTACTACGACCCCAGCAAGGACCTGTGGCCGAGATCCAGAAGCAGCAGGGCCACGACC 1666
         | ||||||||||||| || | ||||||||| | ||||| ||  ||| |  ||||| |  | ||
Db  3476 GAGTGTACTACGACCCCAAGCAAGGACCTGATCGCCAGAGATCGGAGATCCAGAAGCAAGGCC 3535

Qy  1667 AGTGCGACCTACCAGATCTACCAGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1726
         | || |||||||| |  |||||| |  |||||||||||||||||||| | ||||||||| |
Db  3536 AATGACCTACCAAAATCTACCAGGAGCCCTTCAAGAACCTGAAGACCAGCAAGTACGCAA 3595

Qy  1727 AGATGCGCCACCGCCCACCAACGACGTGAAGCAGCTGAAGCACGTGAAGGCAGTGCAGAAGATCG 1786
         ||||| |     || ||| |||||| ||||||||| | |   |||  |||| |||||||||||||  
Db  3596 GGATGAGGGGTGCCCACACCAACGATGTGAAGCAGCAGCTGCAGAGGCAGTGCAGAAGATCA 3655

Qy  1787 CCATCGAGAGCAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGG 1846
         ||| ||||||||||||||||||||||||||||| | ||||||| |  | |||||| ||||||  
Db  3656 CCACAGAGAGCAGCATCGTGATCTGGGCAAGACTCCCAAGTTCAAGTCAGCTGCCCATACAGAGG 3715

Qy  1847 AGACCTGGGAAGCAGACCTGGTGACTACTGGCAGGCCACCACTGGGATCCCCGAGTGGGACT 1906
         ||||  |||||| | |  ||||||| |  ||||| |||| | ||| |||| | ||||||| | 
Db  3716 AGACATGGGAGACATGGTGACCGAGTACTGGCAAGCCACCTGGATCCCTGAGTGGGGAGT 3775

Qy  1907 TCGTGAACACCCCCCCCTGTGAAGCTGTGTTGTGTACCAGCTGGAGAAGAGAGCCCATCATCG 1966
         |||||||||||||||| ||||||| |||| | |  |||||||||||||||   ||||||| | 
Db  3776 TCGTGAACACCCCCCCTCCTGTGAAACTGTGGTATCAGCTGGAGAAGAAGGAACCCATCGTGG 3835

Qy  1967 GCGCCGAGACCTTCTACGTGGACGGCGGCGCCACCGGCGAGATCGGAGAAGATCGGCAAGGCG 2026
         | || |||||||||||||||||| || |||  || || ||| | |   |||  |||||| | 
Db  3836 GAGCAGAGACCTTCTACGTGGATGGGGCAGCAACAGGGGAGACCAAGCTGGGCAAGGCAG 3895

Qy  2027 GCTACGTGACCGACCGGGGCCGCGCAGAAGATCGTGAGCGAAAGTGGTCACCCTGACCGACACCACCAACCAGA 2086
         ||||||||||||||||  | ||  |||||| | ||| ||  ||| |||| |||||| ||||| ||||||||||||  
Db  3896 GCTACGTGACCGACCGAGGACGAGCAGAAGAGTGGTGACCCTGACTGACACACCAACCAGA 3955

Qy  2087 AGACCGAGCTGCAGGCCATCGCAGGCTCGGCGCAAGATCCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCG 2146
         |||| ||||||||||||| |  | || | |||||| || |||  || | |||||  || |||||||||||||  
Db  3956 AGACTGAGCTGCAAGCCATCTACCTAGCTGTCTGCAAGACCAGCAGGCGGACTGAAGTGAACATCG 4015

Qy  2147 TGACCGACAGCCAGCCAGTACGCCCCTGGCATCATCCAGGCCCAGCCCGACAAGAGCCCGAGAGCG 2206
         ||| ||| | ||||  ||||| | ||||||||||||||  || |||| ||||| |||   || | || 
Db  4016 TGACAGACTCACAGTACGCACTGGGCATCATCCAAGCACAACCAGACCAATCCGAGTCAG 4075

Qy  2207 AGCTGGTGAACCAGCCAGATCATCGAGCAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG 2266
         ||||||||||||||  |||||||||||||| ||||||||||||||||| | ||||||||| | |||| 
Db  4076 AGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACCTGGCATGGG 4135

Qy  2267 TGCCCCCGCCACAAGAGGGCCATCGGCGATCGGCGGCAACGAGCAGATCGACAAGCTGGTGCAGCAAGGGCA 2326
         | ||  | ||||||  |||  | ||| ||  |  || ||  || ||| | |||||| ||||| | || | |  
Db  4136 TACCAGCACACAAAGGAATTGGAGAAAATGAACAAGCTAGATAAATTAGTAGTGCAGTGCTGGGA 4195

Qy  2327 TCCGGAAGGTGCTGTTCCTGACGACCATCGAT 2358
         |||||||||||||||||||||||||| |||||
Db  4196 TCCGGAAGGTGCTGTTCCTGACGACGGGGATCGAT 4227
```

Query Match   48.0%; Score 1178.4; DB 4; Length 9010;
Best Local Similarity   69.7%; Pred. No. 1.2e-188;
Matches 1642; Conservative   0; Mismatches 701; Indels 13; Gaps   3;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCGCAGCAACTTCA   73
               |||||||| ||||| |||||||||||||| ||||| |||||||||||||||| || ||| |||
Db   1221 TGGCTGAGGCAATGAGCCAAACAAATAGTGTAAACATACTGATGCAGAGAAAAGCAATTTTA   1280

Qy   74 AGGGCCCCAAGCCGCATCATCAACGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCA   133
               || ||||||||||| |||||| || |||| ||| || |||||||||| |||| || ||
Db   1281 AAGGAAATAAAAGAATGGTTAAAGTGTTTAACTGTGGCAAGGAAGGGCACATAGCCAGAA   1340

Qy   134 ACTGCCGCCCCCCCGCCCCCTGGGAGTGTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA   193
               | ||||| |||||| |||||| || |||||||| | ||| |||||||||| || ||||
Db   1341 ATTGCAGGGCCCCTAGGAAAAAAGGGCTGTTGGAAATGTGGAAATGTGGAAATGTGGGACA   1400

Qy   194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA   253
               || || || ||||| || ||||| |||| | |||||||| |||||||| ||||||| |
Db   3004 GGAGTTCGTGAACAACACCCCCTCCTTGTTGAAACTGTTGAAACTGTATCAGGTATCAGGAACCAT   3063

```
Db  1401 AAGACTGTACTGAGAGGCAGGCTAA-TTTTTAGGGAAAATTTGGCCTTCCCACAAGGGA 1459
Qy   254 AGGCCCGCGAGTTCCCAGGCAGAGAACCGCCCAACAGCCCCACCACGCGCGAGCTGC 313
         ||||||  ||||||||||||  || |||  ||||||||||||| |||  |||||||||
Db  1460 AGGCCAGGGAATTTCCTTCAGAACAGGCCAGAGCCAACAGCCCCACCACGAGAGAGCTTC 1519

Qy   314 AGGTGCGCGGCGACAACCCCGCGAGCGAGGCCGGCGCCGAGCGGGCCAGGGCACCCTGACT 373
         ||| |||||||||||||||||||||  ||||||||||||||||  |||||||||||
Db  1520 AGGTTCGAGGAGACAACACCCGCTCCGAGGAGCAGGAGTCGAAAGACAGGGAAGCCTTAACT 1579

Qy   374 TCCCCCAGATCACCCTGTGACCAGCC... 433
Db  1580 TCCCTCAAATCACTCTTTGCCAGCGACCCCGCCGACACCCTGTCTGCGAGAGATCAAGTAAGGAGTAGGGGGCAAATA 1639

Qy   434 AGGAGGCCCTGCTGGACACCCGCGCGACACCTGCTCGGAGGA... 493
Db  1640 AGGAGGCTCTCTTAGGGGAGCCAAGAAAAACAAGAATTGATACATGGTTTATCGAAGTAGACAATATG 1699

Qy   494 GCAAATGGAAGCCCAAGATCGGCGCATCGGCGAGAGGCCATCGGAGAGATGAAAATGAT 553
Db  1700 GCAAATTGGAAACCAAAAAGATGAGGAGGAATTGGAGCGCTTTATCGAAGTAGACAATATG 1759

Qy   554 ACCAGATCCTGATCGGAGATCGGCGCGAGAGGCCATCGGCGCAAGAGGCTATAGGTATCGCCCCCA 613
Db  1760 ATCAAATACCTATGCAATCAGTGGTCAGTATATTAGTAGGACCTA 1819

Qy   614 CCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGCTGCACCTGAACTTCC 673
Db  1820 CACCTTGTCAACATAATGTTGACTCAGCTTGGATGCAACTAAATTTTC 1879

Qy   674 CCATCAGCCCCATCGAAGCTGCCCGTGAAGCCGTGACCGGCCCCAAGG 733
Db  1880 ACACTAATGATGTAAAACAGTTAACAGAGGCCAGTGCCCTGGGAGACCT 1939

Qy   734 TGAAGCAGTGGCCCCTGACCGAGGAGATCAAGGCCTGCGAGGAGA 793
Db  1940 TTAAACAATGGCCATTGCAGTACTCAGCAGCAATTTGTGAAGAAA 1999

Qy   794 TGGAGAAGGAGGGCAAGAATACCAAGATCGGCCCGAGAACCCCTACAACCCCGTGT 853
Db  2000 TGGGAGAAGGAAGGCAAAAATTACAAAATTCAATACTCCAGTAT 2059

Qy   854 TCGGCCATCAAGAGAGGCAGCACGACCAAGTGGCGCAAGCTGGCGCAAGTCGGTGGAGCTGCGAGCTGA 913
Db  2060 TTGCCATAAAAAGAAGGACAGTACTAAGTGCGAAAATTAGTAGATTTCAGGGAACTCA 2119

Qy   914 ACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGAGTTCAATTAGAAGGTATCCCCGCGGCCTGA 973
Db  2120 ATAAAAGAACTCAAGACTTTGCGGAGAACTTTGGGAAGTTGCTTTTCGGAAGTTTCCTTTAA 2179

Qy   974 AGAAGAAGAAGAGCCGTGACCGTGCTGACCGTGCTGACAGTGTCGACGTGGGCGACGTGCTGACGTGGGCCCTTGG 1033
Db  2180 AAAAGAAAAAATCAGTGACAGTACTGACTTTAGAGCCCTTTAGAGCCTTCAGATTTCTTAG 2239

Qy  1034 ACCAGGACTTCCGCAAGCAGTACACCGGCCTTCACCATCCCAGCATCAACAACGAGACCCCG 1093
Db  2240 ATGAAAGCTTCAGGAAGTACATATACTGGAAATTGCGAAATTCACCAATTCACCAAACACCAG 2299

Qy  1094 GCATCCGCTACCAGTACAGACCTGTCGATGTGATCAATATAATGTGCTTCGAGGCTTGGAGGCTCGACTTAGGAATTCC 1153
Db  2300 GGATTAGATATCAATATAATGTCATTGGAGCGTGGGACATGGAAAGGAATGCACCAGCAATATTCC 2359

Qy  1154 AGATCGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCGCACAAAATCGTGATCT 1213
Db  2360 AGAGTAGCATGAAGATTGACCTTTAGAGCACTTGGGGATGTGGGATGCATATTTTTCAGTTCATCT 2419

Qy  1214 ACCA------GGCCCCCCGTACGTGGGCAGCAGCACCTGGACATCGGCCGCACCGCGCCA 1267
Db  2420 ATCAATATAATGGATGAATGTGCTTCGAGGATCTGATGTATTAGGAATGACACATAGAGCAA 2479

Qy  1268 AGATCGGCGAGAAGCTGCGCAAGCACCTGACCTGTGCTGGGCGTGGGGATTTAAAGTGGGGATTTAC 1327
Db  2480 AAATAGAAGAGTTAAGAGAACATCTATTAAAGTGGGGATTTACCACACCACCAGAGGAAC 2539
```

Right column:

```
Db  1328 ACCAGAAGGAGCCCCCTTCGTGCCCAT------CGAGCTGCACCCGACAAGTCGGACCG 1381
Qy  2540 ATCAGAAAGAACCCCCCATTTCTTTGATGGGTATGGGTATGAACTCATCCTGACAAATGGACCAG 2599

Db  1382 TGCAGCCCATCGAGCTGCCCGAGCAAGAAGGAGAGCTGGACCTGGACTGTTGGACTGGAGAAGCTGG 1441
Qy  2600 TACAGCCTATACAGCTGAAATCGGACTGGACTGTTAATGATATACAGAAGAGTTAG 2659

Db  1442 TGGGCAAGCTGAACTGGGCCAGCCAGATCTCACCCGGCATCAAGGTGCGCCAGCTGTGCA 1501
Qy  2660 TGGGAAAATTAAAACTGGGCAAGTCAAGTTACGCAGGGATTAAAGTAAGGCAACTTTGTA 2719

Db  1502 AGCTGCTGCGGCGGCCAAGGCCCTGACCGTGCCCCTGACCGGAGGAGGCCGAGGC 1561
Qy  2720 AACTTCCTTAGGGGAGCCAAGACTAACAGTACCACTAGTACACAGACAGTAACAGAAT 2779

Db  1562 TGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGGAGCCCGTGACGCGGCCCGGTGTACTACGACC 1621
Qy  2780 TAGAATTGCAGAGAACAAGGAAATTTTAAAGAACCAGTACATGGGGTATATTATGACC 2839

Db  1622 CCAGCAAGGAACTTGGTGCCGAGATCCAGAAGCAGGGCCACGACCAGGCATGGACCTACCAGA 1681
Qy  2840 CATCAAAAGACTTGATAGCTGAAATACAGAAACAAGGGCATGGACCAATATCAAA 2899

Db  1682 TCTACCAGGAGCCCTTCAAGAACCTGAGAGCCCGCAAGTACGCCGCAAGATGCGCACCCCGCCC 1741
Qy  2900 TTTACCAGGAACCATTCAAAAATCTGAAAAATCTGAAAAATGAGGAGCAGCCC 2959

Db  1742 ACACCAACGACGTGAAGCAGCTGACCCGAGGCCGTGCAGAAGATCGCAGAAGATCGGAGAGCATCG 1801
Qy  2960 ACACTAATGATGTAAAACAGTTAACAGAGGCCAGTGCCAAAAAATAGCCCTGGAGCATAG 3019

Db  1802 TGATCTGGGGCAAGACCCCCAAGTTCCGGCTGCCCATCGAGAGGAGACCTGGGAGACCT 1861
Qy  3020 TAATATGGGGCAAGATTCCTAAAATTTAGACTACCCATTCAAAAAGAAAACATGGGAAACAT 3079

Db  1862 GGTGGACCGACTACTGGCAGGAGCACCTGGACTGTGGGTACCAGCTGAGAAGGAGCACACCCCCC 1921
Qy  3080 GGTGACAGCAGACTATTGGCAAGCCACCTGGATTCCTGAGTTGGGAGTTGTTAATACCCCTC 3139

Db  1922 CCCTGGTGAAGCTGTGTGTACCAGCTGGACTGTGGAGAAGGAGCGCCATCATCGGGCGCGCCAACCTTCT 1981
Qy  3140 TCTTAGTAAATTATGGTACCAGCTGGACAAAGAACCATAGTAGGGAAAGAAAACCTTCT 3199

Db  1982 ACGTGACGCGCCGCCAACCCAGGACTTCTGGGAGGTGCACTGGAGTTCAATTAGAAGTTCAATTAGAAGACA 2041
Qy  3200 ATGTAGATGGAGCAGTTAGGTAGGGAAACTAAATTAGGAGACCAGGGTATATTACTGACA 3259

Db  2042 GGGGCCCGGCAGAAGATCGTGAGCGTGCGGCAGCTGACCGAGACGCTGCAGGAGACCCACCAGG 2101
Qy  3260 GAGGAGGCAAAAGTGTTACTGTAACTGAAACACAAATCAGAGACTGAATTACAAG 3319

Db  2102 CCATCCAGCTGCCCTGCAGGACAGCGGCAGAGTGAGCGGCACATGTGACCGCCAGT 2161
Qy  3320 CAATTTACCTAGCTTTGCAAGATTCAGAATCAGAGATTCAGATCAGATCAGTAACTGACTCACAGT 3379

Db  2162 ACGCCCTGGGCATCATCCAGCCGGCCTGGAAGGGCAGCCCCAGCATCTTCC 2221
Qy  3380 ATGCGGTTAGGAATGCATTCAAGCAACATCCAGATAAGGAGTGAATCAGTAGTTAGTCAACCAAA 3439

Db  2222 TCATCGAGCAGCTGCTGATCGAGCAGAAGAAGAAGGTGTACCTGGCGTGGTGGCCCCAAGG 2281
Qy  3440 TAATAGACAATTAATAGAAGAGGAAAAAGGGTCTACCTGTCATGGGTACCACGACATAAAG 3499

Db  2282 GCATCAGCGGCGCAACAAGAGCGCAGATCGGCAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGT 2341
Qy  3500 GAATTGGAGGTAATGAACAGGTAGATAATAAATTGTTAGTAGACAAGGAATCAGGAAAGTGTGCTGT 3559

Db  2342 TCCTTGGACGGCATCGA 2357
Qy  3560 TTCTAGATGGAATAGA 3575
```

RESULT 12
US-09-184-418C-9
; Sequence 9, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 9
; LENGTH: 8972
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=96ZM751.3; 137..1632:gag; 1419..4435:pol;
; OTHER INFORMATION: 4380..4958:vif; 4898..5188:vpr; 5169..7814:tat;
; OTHER INFORMATION: 5308..7938:rev; 5407..5667:vpu; 5585..8128:env;
; OTHER INFORMATION: 8130..8753:nef
US-09-184-418C-9

  Query Match   46.7%;   Score 1147.4; DB 4;   Length 8972;
  Best Local Similarity   69.3%;   Pred. No. 1.8e-183;
  Matches 1647; Conservative   0; Mismatches 696;   Indels 32;   Gaps   5;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCAGGCCAACATCCTGATGCAGCGACCAACTTCA 73
          ||||| || || |||||| ||||| |||||| || ||||| || || || |||
Db   1214 TGGCTGAAGCAATGAGCCAAGTAACAATACAAACATAATAATGATGCAGAAAAGCAATTTTA 1273

Qy   74 AGGGCCCAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCCCA 133
          ||||||||||| | || |||||| |||| ||||| |||||||| ||||| |||||| ||
Db   1333 GATCACCATCAATATTCCAGATGCAGTAGCATGATAGAAAATTCTTAGGCCCTGTCA

Qy   1274 AAGGCCCTAAAAGAATTGTTAAATGTTCAACTGTGGCAGGGAAGGGCCATATAGCCAGGA 1333
          ||||||||| || ||| |||||| ||||||||| |||| ||||||||||| ||||
Db   1333 GATCACCATCAATATTCCAGATGCAGTAGCATGATAGAAAATTCTTAGGCCCTGTCA

Qy   134 ACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
          ||| || || ||||| || || ||||| ||||| |||| || || || || ||
Db   1393 ACCAGGAAATAGTTATCTATCAATATGAGGAGCTGCCAAGCACCTGCTGCGTGGGGCTTCA

Qy   1334 ATTGCAGGGGCTCCTGGGAAAAAGAAGGCTGTTGGAAAATGTGAAAGGGCCACCAAATGA 1393
          ||||||||||| ||||| |||| ||| || ||||| ||||| ||||||||||||||
Db   1393 ACCAGGAAATAGTTATCTATCAATATGAGGAGCTGCCAAGCACCTGCTGCGTGGGGCTTCA

Qy   194 AGGACTGCACCGAGGGGCCAGGCCAACTTCTTCCCCGAGGACCTGCCTTCCCCAGGGCA 253
          ||||| ||| |||||||||| ||||| |||||| ||||||| || |||| ||||| ||
Db   2473 TAGGGCAACACAGAGCAAAAATAGAGGAGTTAAGAGACAACACCTATTGAGATGGGGATTTA

Qy   1394 AAGACTGTACTGAGAGACAGCTAA-TTTTTAGGGAAAAATTTGCCTTCCCAGAAGGGG 1452
          |||||| || |||||| |||| ||| ||| |||| |||| ||||||||
Db   2533 CTACACGAGCAAGAAGCATCAGAAGAGAAAAGAGCCCCATTTCTTTGGATGGGGTATGAACTCC

Qy   254 AGGCCCGGAGTTCCCAGCGGAGCAG----------AACCGCGCCAACAGCC 295
          |||| |||||| ||||| |||| ||||| ||||||
Db   1512 AGGCCGGGAACTTCCTTCAGAACAGACCAGAGCCAACAGCCCCACCAGCTCCAACAGCC

Qy   1453 AGGCCGGGAACTTCCTTCAGAACAGACCAGAGCCAACAGCCCCACCAGCTCCAACAGCC 1512
          |||| ||||| ||||| |||| || |||| | |||| ||||||||| | ||||||
Db   2593 ATCCTGACACAATGACAGAGTACAGCAGCCTATAAAGTGTGACTGTCA

Qy   296 CCACCAGCCGCGAGTCGCAGGTGCGCAGGTCGACAACCCCAGCGGAGGCAGGCCGAGC 355
          ||||| || |||||||||||| | || || || ||| || ||||| ||||| |||
Db   2652 ATCCTGACACAATGACAGAGTACAGCAGCCTATAAAGTGTGACTGTCA

Qy   1513 CCACCAGCAGAGAGCTTCAGGTTCAGGAGGAGACAACAACCCCTGCCCCAGCAGCAGAAA 1572
          ||||| ||||||||||||||| ||||| ||||||||||| |||||| ||||||||||
Db   1572 ACGACATCCAGAGAGCTGAACTGGGCGCAAGTGAACTCACCCGGCAGTCACCCGGCATCA

Qy   356 GCCAGGCACCCGAACTTCCCCCAGATCACCCCTGTGGCAGCGCCCCCTGGTGAGCATCA 415
          ||||||| || |||||||||| |||||||| || || ||||| |||||||| || |||
Db   1424 ACGACATCCAGAGAGCTGAACTGGGCGCAAGTGAACTCACCCGGCAGTCACCCGGCATCA 1483

Qy   1573 GACAAGGAACCCTAACTGCCCTCAAATCACTCTTTGGCAGCGACCCCTTGTCTCAATAA 1632
          || ||| ||||| ||| |||| ||| | || |||||| || |||| | || ||||
Db   2653 ATGATATACAGAAGTTAGTGGGAAAATTAAACT-GGCAAGTGCAGGCCCTCAAGCACCTA

Qy   416 AGGTGGGCGGCCAGATCAAGGAGCCCCTGCTGGACACCCGGCGGCCGCAACACCGTGCTGG 475
          ||||| || || || ||||||| ||| |||| | || || || |||| | || |||
Db   1543 AGGTGCGCCAGCTGCTGCAAGTGCTGCAAGTGCTCGCCAAAGCACAACAGACAGTACCAT

Qy   1633 AAGTAGGGGGTCAGATGAGATAAAGGAGGAGCTCTTCTTGGATACAGAGCAGATATTAG 1692
          ||||| || || |||||||||||| |||||||| || | || ||||| |||| |||||
Db   2771 AGGTGCGCCAGCTGCTGCAAGTGCTGCAAGTGCTCGCCAAAGCACAACAGACAGTACCAT

Qy   476 AGGAGATGAGCGGCTGCCCAGAAAATGGAAACCAAAAATGATAGAGGAATTGGAGGTTTTA 535
          |||||||| | ||||||| ||| || || |||| || || ||||| || ||||
Db   1603 ACGGCGTGTACTACGACCCCAGCAAGAAGCCACACG

Qy   536 TCAAGGTGCGCAGTACGACCCAGATCGAGATCTGAGATCTGAGAAGGGCAAGAAGGGCATCGCCGCA 595
          |||||||| | | || || || ||||| ||||||||||||||| || | ||||
Db   1663 ACGGCGTGTACTACGACCCCAGCAAGAAGCCACACG

Qy   1753 TCAAGTAAGACAGTATGATCAAATACTTATAGAAATTTGTGACGCCACCATTCAAAAAAAAGGCTTAGGTA 1812
          ||||| ||| ||||||| ||||| |||| || ||| ||||| | || || ||| | || | ||
Db   1812 ACCAGTGGACCTACCAGATCTCAAGATCATCGGCCCAACATGCTGCCGACCCAGCTGG

Qy   596 CCGTTGCTGATCGGCCCACCCCCGTGAACATCATCGGCCCAACATGCTGCCGACCCAGCTGG 655
          || ||||| || || ||||| ||||||| ||||||||||| || ||||||||||| ||||
Db   1664 ACCAGTGGACCTACCAGATCTCAAGATCATCGGCCCAACATGCTGCCGACCCAGCTGG 1723

Db   1813 CAGTATTAGTAGGACCTACACCTGTCAACATAATTAATTGGAGAAATATGTTGACCCAGCTTG 1872

```
Qy    14 TGGCCGAGGCCATGAGCCAGGCCACCAGGCGCCACCAGGCCCTGCAGGAAGA   73
            |||||||  ||| |||| | |   |||||||  || ||||||||||||
Db  1216 TGGCTGAGGCAATGAGCCAATCACATAG---TAACATAATGATGCAGAGGAGGCAATTTTA 1272

Qy    74 AGGGCCCAAGCGCGATCATCAAGTGCTTCAACTGCGGCAAGGCACATGCCGCCGCA  133
            |||||| |||||||||| |||||||| ||| ||| || |||| || ||||||||||
Db  1273 AAGGGCCTAAAAGAAATTGTTAAATGCTTCAACTGTGGCAAGGGACACATAGCCAGAA 1332

Qy   134 ACTGGCCGCCCCCGAGAACTGGTGGACAACTACTGGCAGGCCACCTGGATCCCGAGTGGG  193
            | || | ||| || ||||||| |||||||| ||  ||||||||||||||||||
Db  1333 ATTGCAGGGCCCCTAGAAAAAGAGGCTGTTGGAAATGTGGCAAGGAAGGACACCAAAATGA 1392

Qy   194 AGGACTGCAACACCCCCCCTGTGGTACCAGCTGAGCAAGGAGGAGCCCATCA  253
            ||||| |||||||||   |||||| ||||| ||| ||||||| || ||||
Db  1393 AAGACTGTACTGAGAGACAGGCTAA-TTTTTAGGCCAGCTTGGCCTTCCCACAAGGGG 1451

Qy   254 AGGCCGCCGAGTTCCCAGCGGCGCCGCCAACCGCGGCCAGAACGGCCAAGG  313
            |||| ||||| |  |||   ||| | | |||||||| | || || |||
Db  1452 AGGCCAGGGAATTTCCTTCAAAACAGGCCAGCCAACAGCCCCACCAGAGAGCTTC 1511

Qy   314 AGGTGCGGCGACAACCCCCGCGAGCGCGGCCGCAGCGCCACCCTGAACT  373
            ||| ||||| |  |  |||| |||| |||| | || ||||| |||||
Db  1512 AGGTTCAAGGAGGACACACCCCCGCTCCGAAGCAGGGAACCCCTTAACT 1571

Qy   374 TCCCCCAGATCACCCTGTGCAGCCCCCTGGTGAGCATCAAGTGGGGCGGCCAGATCA  433
            |||| |||| ||| ||| |||| || ||||| |||||||| | |||| |||||
Db  1572 TCCCTCAAATCACTCTTTGGCAGCGACCCCCTTGTCTCAATAAAAGTAGGGGCCCAGATAA 1631

Qy   434 AGGAGGCCCTGCTGGACACCGGCGCCGACCACCGTGCTGGAGGAGATGAGCCTGCCCG  493
            ||| ||| || |||| || ||| || || | | || |||||||| | |||| |||
Db  1632 AGGAAGCTCTCTTAGACACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGCCAG 1691

Qy   494 GCAAGTGGAAGCCAAGATGATCGGCGGCATCGGCGCTTCATCAAGGTGCGCCAGTACG  553
            |||| |  |  || |  ||  ||||| || || |||  ||||| | | | ||||
Db  1692 GAAGATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATG 1751

Qy   554 ACCAGATCCTGATCGAGAGGCCATCGGCCAACCGTGCTGATCGGCCCCA  613
            | || ||| | || || ||||  | |  ||||  | |||| ||||
Db  1752 ATCAAATACTTATAGAAATTTGTGGAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTA 1811

Qy   614 CCCCCGTGAACACATCGAGCACCTCGGCCGCAACATGCTGACCCAGCTGGCCTGCACCCTGACCTTCC  673
            |  ||| ||||||   |||| || | ||  || |   ||| |   ||| ||||
Db  1812 CACCTGTCAACATAATTGGAAGAAATATGTTGACTCAGCTTGGATGCACTTTAAATTTTC 1871

Qy   674 CCATCAGCCCCATCGAGCACCGTGCCCGTGAAGCTGAAGCTGAAGCGCCGGCCGCATGGACGGCCCCAAGG  733
            ||||| |  || | ||   || |   | |||  | |||| || |||| ||| |||||||||
Db  1872 CAATTAGCCCCATTGAAACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAGG 1931

Qy   734 CCATCAGCCCTGACCGGAGAAGATCAAGGCCCTGACCGCGATCACCCCCTGT  793
            | || | || |||| || ||| ||| |  | || | || || |  ||||
Db  1932 TTAAACAGTGGCCATTGACAGAAGAAAAAATAAAAGCATTAACAGAAATTTGTAAAGAAA 1991

Qy   794 TGGAGAAGGAGGGCCAAGATCACCACCAAGATCGGCCCCGAGAACCCCTACACACCCCGTGT  853
            ||||| || ||  ||  || |   || ||   |  |||  || | | |||
Db  1992 TGGAGAAGGAAGGGAAAATTTCAAAAATTGGGCCTGAAAATCCATATAACACTCCAGTAT 2051

Qy   854 TCGCCATCAAGAGACGAGCAGCACCAGTGGCGCAAGTGGCCAAGCTGGTGGAAAATTAGC  913
            | || ||| | || || |||| || ||| |  |   | ||| |||  |||||
Db  2052 TTGCCATAAAGAAAAAGGACAGTACTAAGTGGAGAAAATTAGTAGATTTCAGGGAGCTCA 2111

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCACCCCGCCGGCCTGA  973
             || | || || ||||| ||| | ||| | | |  || ||| |  | || ||| ||
Db  2112 ATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAATACCACACCCAGCAGGTTAA 2171

Qy   974 AGAAGAAGAAGCGGTGACCGTGCTGACGTGGCGCACGCTCACTTCAGCGTGCCCCTGG 1033
            |||||||| | ||| | |  || |  ||| ||| | ||  | || |||| ||||||
Db  2172 AAAAGAAAAAATCAGTAACAGTACTGGATGTGGGGGATGCATATTTTTCAGTTCCTTTAG 2231

Qy  1034 ACGAGGACTTCCGCAAGTACACCGGCTTCACCGCCTCCCAGCATCAACAAGCAGACCCCG 1093
            | ||| | || |  ||||   |||  | |  || ||| |||| | | ||||||| ||
Db  2232 ATGAAGGCTTCCGGAAATATACTGCATTCACCATACCTAGTATAAACAATGAACACCAG 2291
```

Qy 1724 CCAAGATGCGCACCGCCCAACGACGTGAAGCAGCAGCTGACCGAGCCGCCCTGCAGGAAGA 1783
           || ||| ||| |||| || | || || | | || || ||| || ||||
Db 2952 CAAAAATGAGGACTGCCCACACTAATGATGTAAAACAGTTAACAGAGGCGGTGCAAAAAA 3011

Qy 1784 TCGCCATGGAGAGACTGGATCGTGATCTGGCGCAAGACCCCCAAGTTCCGCCTCCATCCAGA 1843
           | ||| ||| | ||| ||||| ||||| || || ||| | ||| |||| | ||||
Db 3012 TAGCCATGCATGAAAGCCATAGTAATATGGGGAAAGATTCCTAAATTTCAGGTACCTACCATCAAA 3071

Qy 1844 AGGAGACCTGGGAGACTACTGGCAGTTCGAGGCCACCTGGATCCCGAGTGGG 1903
           | | || ||||| || ||| |||| | || | |||| || ||||||
Db 3072 AAGAAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGG 3131

Qy 1904 AGTTCGTGAACACCCCCCCTGTGTGGTACCAGCTGTGGTACCAGCTGGGTACAGAAGGAGCCCATCA 1963
           |||| || || || ||||| | ||||||| | ||||| || |||| || | ||||
Db 3132 AGTTTGTTAATACCCCCCCTAGTAAAATATGGTACCAGCTGGAGAAAGAACCCATAG 3191

Qy 1964 TCGGCGCCGAGACCTTCACGTGACGCGGGCGCCGCCAACCGCGACCGAGACCACCAAGG 2023
           | | | || | |||| | ||| ||||| ||||| |||| | || |||| | ||||
Db 3192 CAGGAGCAGAAACTTCCTCAAAACAGGCAGCCAATAGGGAAACTAAATTAGGAAAAGCAGGA 3251

Qy 2024 CCGGCTACGTGACCGACCGGGCGCCGAGAACGTGTGAGCCTGACCGAGACCACCAACC 2083
           | || | |||| |||| | | || | || | | | || ||| ||| | ||
Db 3252 CAGGGTATGTTACTGACAGAGGAAGGCAAAAAAGTTGTCTCCCTAACTGACACAACAAATC 3311

Qy 2084 AGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGACGCGGCGAGGTGAACA 2143
           || || || | |||| | || | || || || || |||| | |||| ||
Db 3312 AAAAGACTGAATTACAAGCAATTCAGTTAGCTTTGCAGGATTCAGGATCAGAAGTAAACA 3371

Qy 2144 TCGTGACCGACAGCAGCAGTACGCCCTGGCATCATCCAGGCCGAACAAGAGCCGGAGA 2203
           | || || | ||||| |||| | | ||||||| || | | ||| ||| |||| ||
Db 3372 TAGTAACAGACTCACAGTATGCATTAGGAATCATTCAAGCACAACCAGATAAGAGTGAAT 3431

Qy 2204 GCGAGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAAGGTGTACCTGAGCT 2263
           || ||||| ||||| || | |||| ||| || ||| || || || ||| | ||||
Db 3432 CAGAATTAGTCAATCAAATAATAGAGCAGTTAATAAAAAAGGAAAAGGTTTACCTGTCAT 3491

Qy 2264 GGGTGCCGCCCAACAGGGCATCGGCGGCAACGACGCAGATCGACAAGCTGGTGAGCAAGG 2323
           ||| |||| || | || | ||| | | | | || ||| || || || |||| |||||
Db 3492 GGGTACCAGCACAAAAGGAAATTGGAGGAAATGAACAAGTAGATAAATTGGTAAGTAGTG 3551

Qy 2324 GCATCCGCCAAGGTGCTGTTTCCTGGACGGCATCGAT 2358
           ||| | | |||| ||||||| ||||| || || ||
Db 3552 GAATCAGGAAAGTGCTGTTTCCAGATGGAATAGAT 3586

RESULT 13
US-09-184-418C-11
; Sequence 11, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 11
; LENGTH: 8959
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=94IN476.104; 138.1613:"gag";
; OTHER INFORMATION: 1418.4428:"pol"; 4361.4939:"vif"; 4879.5169:"vpr";
; OTHER INFORMATION: 5150.7782:tat"; 5288.7933:"rev"; 5378.5638:"vpu";
; OTHER INFORMATION: 5556.8129:"env"; 8131.8754:nef
US-09-184-418C-11

Query Match   46.4%;   Score 1140.2; DB 4; Length 8959;
Best Local Similarity   69.2%;   Pred. No. 2.9e-182;
Matches 1632; Conservative   0; Mismatches 708; Indels   17; Gaps   5;

RESULT 14
US-09-872-733A-3

; Sequence 3, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
;   APPLICANT: The Government of the United States of America, as
;   TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
;   TITLE OF INVENTION: SIV ENV GENES
;   FILE REFERNCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
;   CURRENT APPLICATION NUMBER: US/09/872,733A
;   CURRENT FILING DATE: 2001-06-01
;   PRIOR APPLICATION NUMBER: PCT/US00/34985
;   PRIOR FILING DATE: 2000-12-22
;   PRIOR APPLICATION NUMBER: 60/173,036
;   PRIOR FILING DATE: 1999-12-23
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2467
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
;   OTHER INFORMATION: Immunodeficiency Virus - 1 Pol gene
US-09-872-733A-3

  Query Match   46.3%; Score 1137; DB 4; Length 2467;
  Best Local Similarity   83.4%; Pred. No. 8.6e-182;
  Matches 1319; Mismatches 250; Conservative   0; Indels 12; Gaps 2;

Qy   2278 AGGGGCATCGGCGGCGGCAACGAGCAGCAGAGATCGACAGCAGAGCTGGTGAGCAAGGGCATCCGCAAGGTG 2337

Db   1507 AAAGGAATTGGAGGAAATGAACAAGTAGACGTAGATAAATTAGTCAGTGCTGGGATCCGGAAGGTG 1566

Qy   2338 CTGTTCCTGGACGGCATCGAT 2358

Db   1567 CTGTTCCTGGACGGCGATCGAT 1587

RESULT 15
US-09-184-418C-4
; Sequence 4, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 4
; LENGTH: 8992
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate=92RW009; 139..1624:gag; 1690..4428:pol(N-terminus uncertain
; OTHER INFORMATION: 4373..4951:vif; 4891.5181:vpr; 5162.7801:tat; 5301.7958:rev;
; OTHER INFORMATION: 5403..5648:vpu; 5566.8148:env; 8150.8773:nef;
US-09-184-418C-4

  Query Match   45.5%; Score 1116.8; DB 4; Length 8992;
  Best Local Similarity 68.3%; Pred. No. 2.4e-178;
  Matches 1612; Conservative 0; Mismatches 732; Indels 16; Gaps 4;

Qy   14 TGGGCGAGCCATGAGCGCAGGCCACCAGCGCCAACATCCTGATGCAGCGAGCAACTTCA 73

Db   1221 TGGCTGAAGCAATGAGCCAAGTACAACACCAAACATAAATGATGCAGAGAGGCAATTTTA 1280

Qy   74 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCA 133

Db   1281 AGGGGCCAGAAGAGATTATTAAGTGTTTCAACTGTGGCAAAGAGAGGACACCTAGCCAGAA 1340

Qy   134 ACTGCGCGCCCCCCCGCAAGAAGGCGTGCTGGAAGTGGGCAAGAGGCCACCAGATGA 193

Db   1341 ATTGCAGGGCCCCTAGAAAAAAGGGCTGTTGGAAATGTGGAAAGGAGGACACCAAATGA 1400

Qy   194 AGGACTGCGAGGCCAGGCCAACTTCTTCCCGGAGAAGATGCCGCCAGACC 253

Db   1401 AAGACTGCACTGAGAGACAGGCTAA-TTTTTTAGGGAAAATTGGCCTTCCARCAAGGGGG 1459

Qy   254 AGGCCCGAGTTCCCGAGCGAGAACCGGCCAACAGCCCCACCAGGCGCGGAGCT-- 311

Db   1460 AGGCCAGGAGAATTTTCCCAGAGCAGACTGGAGCCAACAGCCCCACCAGCAGAGAACTTT 1519

Qy   312 -GCAGGTGCGCGGCGGAAGAGATCGGCCGGAAGCGGGCGGCCACCCTGA 370

Db   1520 GGAATGGGGAAGAGATAGCCTCTCCTCTGAAACAGGAGCAGAAAGACAGGGAACCTTTA 1579

Qy   371 ACTTCCCCGATCACCTGTGGCAAGCGGCCCCTCGTGGAGCGCGCCCCTGTGACCGACAGC 430

Db   1580 ATTTCCCTCAAATCACTCTTTGGCAACGACCCCTTGTCACAGTAAAATAGGAGTCAGC 1639

Qy   431 TCAAGGAGGGCCCTGCCTGGAAGATCGGCAGCGAGCGGCAGGAGATGAGCCTGC 490

Db   1640 TAAGAGAGCTCTATTAGATACAGGAGCAGATGATGATACAGTATTAGAAGAAATAAATTTGC 1699

Qy   491 CCGGCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGT 550

Db   1700 CAGGAAAATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAGGTAAAACAGT 1759

Qy   1210 ATCTACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGCGCCACCGC 1263

Db   427 ATCTATCAGTACATGGACCCTCACGTAGGGAATGACCTGGAGGATGACGGCGCACAGG 486

Qy   1264 GCCAAGATCGAGAGAGCTGCGCCAAGGACGTCGGCGGGCTTCACCACCCCCGACAAG 1323

Db   487 ACCAAGATCGAGAGAGCTGAGACTGAGCTGTTGAGGTGGGACTGGGACCACCAGACACAG 546

Qy   1324 AAGCACCAGAAGAGAGCCCCTTCCTGCCCAT------CGAGCTGCACCCGACAAGTGG 1377

Db   547 AAGCACCAGAAGGAACCTCCCTTCCTTCCTGTGGATGGGCTACCAACTTGCATCCTGACCAAGTGG 606

Qy   1378 ACCGTGCGAGCTCAGCTGCTGCCCGAGAAGAGCTGGACCGTGAACGACATCCAGAG 1437

Db   607 ACAGTGCAGCCCATCCGTGCTCGCCTGAGAAGGAAGGACACAGTGTGAACGACATACAGAGAG 666

Qy   1438 CTGGTGTGGGCAAGCCAGATCGGCCAGCCAGATCTACCCCGGCATCAAGTGCGCCCAGTG 1497

Db   667 CTCGTGGGGCAAGTTGAACTGGGCGGGCGGCCAGATCTACCAGCATCAAAGTTAGGCAGCTG 726

Qy   1498 TGCAAGCTGCTGCGCGGGCGGCCAAGGCCCTGACCGACATCGTCGTCCCCTGACCGAGGAGGCC 1557

Db   727 TGCAAGCTGCTTGCGGGAACCAAGGCACTGACACAAGATGAATCCCACTGACAGCAGAGGAAGCA 786

Qy   1558 GAGCTGGAGCTGGCCCGAGAACCGCCGAGATCCTGCGCGGAGCCCGTGCACGGCGTGTACTAC 1617

Db   787 GAGCTAGAACTGGCAGGAGGAACCGCCGAGATCCTGAAGGAGGCCAGTACAATGGAGTGTACTAC 846

Qy   1618 GACCCCAGCAAGGACCTGGTGCCCGAGATCCAGAAGCAGGGCCACGACCAGTGACCTAC 1677

Db   847 GACCCAAGCAAGGACCTGATCCCAGAGATCCCAGAGATCCCAAGGGCCAATGACCTAC 906

Qy   1678 CAGATCCACCAGGAGCCCTTCAAGAACCTGAAGATCGGCAAGTACGCCCAAGATGCGCACC 1737

Db   907 CAAATCCACCAGGAGCCCTTCAAGAACCTGAAGATCGGCAAGTACGCAAGGATGAGGGGT 966

Qy   1738 GCCCACCACCAACGACGTGAAGCAGCTGACCGAGCGCGTGCAGAAGATCGCCCATGGAGAG 1797

Db   967 GCCCACCACCAACGACGTGAAGCAGCAGTGCAGACATGCTGACAGAAGATCACCACAGAGAG 1026

Qy   1798 ATCGTGGATCTGGGGCAAGAAGACCCCCCAGTTCCCGCTCCGCCATCCAGAAGGAGACCTGGGAG 1857

Db   1027 ATCGTGTGGATCTGGGGCAAGACTCCCAGTTCAAGCTGCCCATACAGAAGGAGACATGGGAG 1086

Qy   1858 ACCTGGTGGACCGACTACTGCAGGCCCACCTGAGATCCCGAGTTCGTGAACACC 1917

Db   1087 ACATGGTGGACCGACTACTGGCAAGCCACCTGAGATCCCGAGTTCGTGAACACC 1146

Qy   1918 CCCCCCTGGTGAAGCTGTGTGTACAGTGGAAGGAAGGAGCCCATCATCGTGGCGCCAGGACC 1977

Db   1147 CTCCCCTTGGTGAAACTGTGGTATCAGCTGGAAGAAGGAACCCATCGTGGGGCAGCAGAGACC 1206

Qy   1978 TTCTACGTGGAACGGCCGCGCGCCAACCGCGGCCAAAGATCGGCCAAGGCCGGCTACGTGACC 2037

Db   1207 TTCTACGTGGATGGGGCAGCGCCAACAGCGGAGCCAGCTGGGCAAGGCCAGGCTACGTGACC 1266

Qy   2038 GACCGGGGGCCGGCAGAAGATCGTGACGAAGCTGACCGAGAGGTGAACATCGTGACCGAGCTG 2097

Db   1267 AACCGAGACGACAGAAGTGGTGACCGACACCGACTGACAGCCAACAGCCAACAAGAAGCTGAAC 1326

Qy   2098 CAGGGCCATCCAGCTGCCCTGACCGACCAGAGGTGAACATCGTGACCGACAGC 2157

Db   1327 CAAGGCCATCATCCTAGGTCTGCAAGACATGGTGAAGTGAACATCGTGACAGAGCTCA 1386

Qy   2158 CAGTACGCCCTGCTGCGGCCATCATCCAGGCCCAAGAGAGCGAGAGGGCTGGTGAAC 2217

Db   1387 CAGTACGGCACTGCTGGCATCATCCAACCAGCACAACCAGCCAGCTCAGTGGGTGGTGAAC 1446

Qy   2218 CAGATCATCGAGCAGCTGATCAAGAGGAAGGTGACCTCAGCGCTCATCAAGGTGCCCCAGT 2277

Db   1447 CAGATCATCGAGCAGCTGATCAAGGCAGCGAGCAGCAGAAAGTGTACCTGGGCATGGGTACCAGCACAC 1506

Qy   2840 ACCCATCAAAAGACTTAATAGCTGAATAGCTTGAAATACAGAAACAGGGCATGAACATGACCAATGACATATC 2899

Db   1679 AGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAGGAGCCCGCAAGTACCGCCAAGATGCGCACCG 1738

Qy   2900 AAATTTACCAAGAACCATTCAAAAATCTGAAAACAGGAAACTATGCAAAAAGGAGGAGACTG 2959

Db   1739 CCCACCACCAACAGCTGAAGCAGCTGACCGAGCTGCGAGAATCGCCATGAGCGAGCA 1798

Qy   2960 CCCACACTAATGACGTAAAACAGTTAACAGTGCAGTGCAAAAGATAGCCATGGAAAGCA 3019

Db   1799 TCGTGATCTGGGGCAAGAGCCCCATCCAAGTTCCGCTGCCCATCCAGAAGGAGACCTGGAGA 1858

Qy   3020 TAGTAATATGGGAAAGACTCCTAAATTTAGATTAGATTCCCATCCAGAAAGAAACATGGGAAA 3079

Db   1859 CCTGGTGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGTGGAGTTCGTGAACACCC 1918

Qy   3080 CATCGTGACAGACATATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTTAATACCC 3139

Db   1919 CCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCCAAGGGCCGAGACCT 1978

Qy   3140 CTCCCCTAGTAAATTATGGTACCAGCTAGAGAAAGAACCCATATTAGGAGCAGAGACTT 3199

Db   1979 TCTACGTGGACGGCGCGCCAAGATCGTGAGCCTGACAGAGGCCAAGATCGGCAAGGGCCGGCTACGTGACCG 2038

Qy   3200 TCTATGTGGATGGAGCAGCTAATCGGGAAAACTAAATAGGAGAAAGCAGGGTATGTTACTG 3259

Db   2039 ACCGGGGCCCGCAGAAGATCGTGAGCCTGACAGGCGTGACCTACTTCAGCGTGCCCC 2098

Qy   3260 ACAGAGAAGGCAGGAAAATTGTTTCTTAACTGAAACATCAGGATAAACAAATCAGAAGACTGAATTAC 3319

Db   2099 AGGGCCATCAGCTCCCTGCAGGACGCGCAGCAGCCGAGGTGAACATCGTGACCGACAGCC 2158

Qy   3320 AAGCAATTCAGCTAGCTTACAGGATTCAGGATCAGAGTCAAGTAAACATAGTAACAGACTCAC 3379

Db   2159 AGTACGCCCTGGGCATCATCCAGGCCCAGGCTGAAGGCCAGCCCAACCAGAGTGTACCGCAGTCGGTGAACC 2218

Qy   3380 AGTATGCATTAGGAATCATTCAAGCACAACCAGATAGCAGCGATCAGTGGAGGCAGTCAATC 3439

Db   2219 AGATCATCGAGCAGCTGATCAACGAAGAGGAGAAGTGTACCTGAGCTGGTGGGCTGCCCGCCCACA 2278

Qy   3440 AAATAATAGAACAGTTAARTAARAAAAAGGAAAGAGTCTACCTCGTCGTCATGGGCACATA 3499

Db   2279 AGGGCCATCGGCGGCAACGAGCAGATCGACACGAGCTGGTGAGCAAGGCATCCGCAAGGTGC 2338

Qy   3500 AAGGAATTGGAGGAAATGAACAACAGTAGATAAATTAGTAGTAAGTGGAATCAGGAGAGTGC 3559

Db   2339 TGTTCCTGGACGGCATCGAT 2358

Qy   3560 TGTTTCTACAGTGGAATAGAT 3579

Search completed: April 10, 2004, 16:22:57
Job time : 129.045 secs

Qy   551 ACGACCAGATCCTGATCGGAGATCTGCGGCAAGAGAGGCCATCGGCACCGTGCTGATCGGCC 610

Db   1760 ATGATCAAATACTTATAGAGAAAATTTGTGAAAAAAAGCTTATAGTAGTACAGTATTAGTAGGAC 1819

Qy   611 CCACCCCGTGAACATCATCGGCGCAACATGCTGGGCTGCACCCTGACT 670

Db   1820 CTACATCTGTCAACATAATTGGAAGAAATATGTTGACCCAGATTGGTGTTACTTTAAACT 1879

Qy   671 TCCCCATCAGCCCATCGAGACCGTGCCCGTGAAGTCGAAGTGAACCGGCCCA 730

Db   1880 TTCCAATTAGTCCTATTGAGACTGTACCAGTAGCATTAAAGCAGGATTGGATGGCCCAA 1939

Qy   731 AGGTGAAGCAGTGCCCCCTGACCGAGGAAGATCAAGGCCCTGACCGGCCATCTGCGAGG 790

Db   1940 AGGTTAAACAATGGCCATTGACAGAGAAAAATAAAGCATTAAGAGAAATTGTACAG 1999

Qy   791 AGATGGAGAAGGAGGCCAAGATCACCAAGATCGGCCCGAGAACCCTACAACACCCG 850

Db   2000 AAATGGAAAAAGAGGGGAAAAATTTCAAAAATCGGCCTGAAAATCCATATAAACTCCAG 2059

Qy   851 TGTTGCCATCAGAGAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGAGTTCCGGAGC 910

Db   2060 TATTTTGCCATAAAAAAGAAAGAACAGACAGTACTAAGTGGAGAAAATTAGTAGATTTCAGGGGAAC 2119

Qy   911 TGAACAAGCGCACCAGGACTTCTGGGAGGTGCAGCTGGACGTGGGCGCACGCCTACTCAGCGGT 970

Db   2120 TCAACAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGGATACCACACCCAGCAGGGT 2179

Qy   971 TGAAGAAGAAGAATCAGTGACACTGTGACCGTGGACGTGGCACGGCGTACCTTCAGCGTCTTCCTT 1030

Db   2180 TAAAGAAGAAAAAATCAGTAACAGTACTGGATGTGGGGGATGCATACTTTCAGTTCCT 2239

Qy   1031 TGGACGAGGACTTCCGCAAGTACACCGCGCTTCACCATCCCAGCATCAACGAGACCC 1090

Db   2240 TAGATGAGAGCTTCAGGAAGTATACTGCATTCACCATACCTAGTATAAACAATGAAACAC 2299

Qy   1091 CCGGCATCCGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGCCAGCCCAACCATCT 1150

Db   2300 CAGGAATTAGGTATCAATATAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATAT 2359

Qy   1151 TCCAGAGCAGCATGACCAAGATCCTGAAGCCTTCGCGCCGCAACCCCGAGATCGTGA 1210

Db   2360 TCCAAAATAGTAGTAGATGAACAAAATCTTAGAGCCCTTTAGGGCACAAAAACCAAGAAATAGTGA 2419

Qy   1211 TCTACCA------GGCCCCCTGTACGTGGGCAGCGCGACCTGGCAGTCGGCCAGCACCGCG 1264

Db   2420 TCTATCATATATGGGATGACTTGTACCGTTAGGATCTGACTTAGAAATAGGGCAACATAGAG 2479

Qy   1265 CCAAGATCGAGGAGCTGCGGCAAGCACCTGCCGCGGGGCTTCACCACCCCGACAAGA 1324

Db   2480 CAAAAATAGAGGAGTTAAGAGAACATCTATTAAAGTGGGGATTTACCCACACCAGACAAGA 2539

Qy   1325 AGCACCAGGAGAAGGAGCCCCCTTCCTGCCCAT-------CGAGCTGCGACCCGACAGTGGA 1378

Db   2540 AACATCAGAAAGAACCCTCCATTTCTTTTGGATGGGTATGAACTTCATCCTGACAAATGGA 2599

Qy   1379 CCGTGCAGCCGCATGAACTGCTGCCCGAGCAGCTGGCCGAGACCTGAACTGTGTCATGTGTCATGAGCTGATATACAGAAGC 1438

Db   2600 CAGTACAACCTATACAGCTGCCCAGAAAAAGAGAGTAGTGGACTGGTGTCATGAGATACAGAAGT 2659

Qy   1439 TGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGT 1498

Db   2660 TAGTGGGAAATTAAAACTGGCCAAGCAAGTTTACCCAGGGGTTAAAGTAAGGCAATTGT 2719

Qy   1499 GCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGAGGCCCTGACCGAGGAGGCCG 1558

Db   2720 GTAAACTCCTTAGGGGAACCAAAGCATTAACAGACATACCACTAACTGAAGAAGCAG 2779

Qy   1559 AGCTGCGAGCTGGCCGAGACCCGAGAACCCGTGCACGGCGTGTACTACG 1618

Db   2780 AATTAGAATTGGGCAGAAAACAGGGAAAATTTAAAAGAACCAGTACATGGAGTATATTATG 2839

Qy   1619 ACCCCAGCAAGGACCTGTGCGCCGAGATCCAGAAGCCAGGGGCCACGACGACCAGTGGACCTACC 1678

OM nucleic - nucleic search, using sw model

```
              GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

Run on:    April 10, 2004, 06:34:42 ; Search time 4138 Seconds
                                      (without alignments)
                                      17774.420 Million cell updates/sec

Title:     US-09-610-313-31
Perfect score: 2463
Sequence:  1 gtcgacgccaccatggccga........gggctagcaccggtgaattc 2463

Scoring table: IDENTITY_NUC
           Gapop 10.0 , Gapext 1.0

Searched:  2751289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:    55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  EST:*
       1:  em_estba:*
       2:  em_esthum:*
       3:  em_estin:*
       4:  em_estmu:*
       5:  em_estov:*
       6:  em_estpl:*
       7:  em_estro:*
       8:  em_htc:*
       9:  gb_est1:*
      10:  gb_est2:*
      11:  gb_htc:*
      12:  gb_est3:*
      13:  gb_est4:*
      14:  gb_est5:*
      15:  em_estfun:*
      16:  em_estcm:*
      17:  em_gss_hum:*
      18:  em_gss_inv:*
      19:  em_gss_pln:*
      20:  em_gss_vrt:*
      21:  em_gss_fun:*
      22:  em_gss_mam:*
      23:  em_gss_mus:*
      24:  em_gss_pro:*
      25:  em_gss_rod:*
      26:  em_gss_phg:*
      27:  em_gss_vrl:*
      28:  gb_gss1:*
      29:  gb_gss2:*
```

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

	Result No.	Score	% Match	Query Length	DB	ID	Description
	1	86.8	3.5	2598	11	AY103647	AY103647 Zea mays
	2	85.6	3.5	951	12	BM321451	BM321451 rockefell
c	3	80.6	3.3	869	14	CK159167	CK159167 FGAS04056
	4	80.2	3.3	1132	12	BM320864	BM320864 rockefell
	5	79.6	3.2	1165	12	BM320900	BM320900 rockefell
	6	79	3.2	867	12	BM321430	BM321430 rockefell
	7	76.6	3.1	1550	12	BM321022	BM321022 rockefell
	8	75	3.0	862	12	BM321023	BM321023 rockefell
	9	74.4	3.0	914	28	BZ568300	BZ568300 pacs2-164
	10	73.2	3.0	853	12	BM321393	BM321393 rockefell
c	11	72.8	2.9	566	12	BM587428	BM587428 170006873
	12	72.6	2.9	788	14	CB643171	CB643171 OSJNEb03L
	13	72.4	2.9	753	29	CC675888	CC675888 OGWCO51TH
	14	72.2	2.9	640	12	BE601575	BE601575 HVSMEh009
	15	72.2	2.9	688	14	CB648640	CB648640 OSJNEb12C
	16	72.2	2.9	764	14	CB651670	CB651670 OSJNEb16L
	17	72.2	2.9	766	14	CB642928	CB642928 OSJNEb03F
	18	72.2	2.9	809	14	CB641397	CB641397 OSJNEb01A
	19	72.2	2.9	841	14	CB651502	CB651502 OSJNEb16H
c	20	72.2	2.9	851	29	CG260623	CG260623 OGWGE95TV
	21	71.8	2.9	1933	11	AY325173	AY325173 Rattus no
	22	71.6	2.9	500	12	BM372120	BM372120 EBro03_SQ
	23	71.6	2.9	538	12	BM368580	BM368580 EBem08_SQ
	24	71.6	2.9	540	9	AJ471121	AJ471121 AJ471121
	25	71.6	2.9	566	13	BQ464692	BQ464692 HF02P20r
	26	71.6	2.9	575	13	BU976068	BU976068 HA03B08r
	27	71.6	2.9	576	13	BU984666	BU984666 HF04K03r
	28	71.6	2.9	578	12	BM377112	BM377112 EBem05_SQ
	29	71.6	2.9	579	10	BF253640	BF253640 HVSMEf000
	30	71.6	2.9	582	12	BM372871	BM372871 EBma04_SQ
	31	71.6	2.9	586	13	BU994273	BU994273 HM06J03r
	32	71.6	2.9	588	13	BQ765498	BQ765498 EBro03_SQ
	33	71.6	2.9	596	13	BU997745	BU997745 HI08O07r
	34	71.6	2.9	600	9	AL508778	AL508778 AL508778
	35	71.6	2.9	615	13	CA009718	CA009718 HU14O18r
	36	71.6	2.9	621	13	BU986998	BU986998 HF13I08r
	37	71.6	2.9	625	13	BQ768295	BQ768295 EBro08_SQ
	38	71.6	2.9	637	13	BU999333	BU999333 HI14D01r
	39	71.6	2.9	641	13	BU978712	BU978712 HA14B02r
	40	71.6	2.9	645	13	BU977527	BU977527 HA11J06r
	41	71.6	2.9	646	13	BU968956	BU968956 HB10B12r
	42	71.6	2.9	655	13	BU996444	BU996444 HM13K04r
	43	71.6	2.9	656	13	BM370222	BM370222 EBro08_SQ
	44	71.6	2.9	659	13	CA031811	CA031811 HX11D06r
	45	71.6	2.9	660	13	BQ762385	BQ762385 EBro01_SQ

ALIGNMENTS

```
RESULT 1
AY103647
LOCUS       AY103647               2598 bp    mRNA    linear    HTC 16-OCT-2002
DEFINITION  Zea mays PCO142084 mRNA sequence.
ACCESSION   AY103647
VERSION     AY103647.1  GI:21206725
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 2598)
  AUTHORS   Rainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 2598)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
```

maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES   Location/Qualifiers
     source   1..2598
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:638378"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
ORIGIN

Query Match   3.5%;   Score 86.8; DB 11; Length 2598;
Best Local Similarity   42.1%;   Pred. No. 0.59;
Matches 813; Conservative   0; Mismatches 1102; Indels   15; Gaps   5;

```
Qy    296 CCACGACCGCGGAGCTGCAGTGCGCGGGACAACCCCGCAGCAGGAGGCGCGGCGAGC  355
          ||| |||| || |||| |||  |||  || |  ||||| |  |   ||||
Db     51 CCACCACCTCCACTTCCTGCACCGGCGCCCAACCAACCACCGGCAGGCGGCGCAA   110

Qy    356 GCCAGGGCACCCGAACTTCCCCAGATCACCCTGTTGGCAGCGCCCCTGTGTGAGCATCA  415
          |||  ||||| |
Db    111 TGGCGGGCAACGGCGCCATCGTGGTGGAGAGCGCCGCTGAACTGGGCGCGGCGCGGCGG  170

Qy    416 AGGTGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGGCTGG   475
          ||||
Db    171 AGCTGGCCGGGAGCCACCTGGACGAGGTGAAGCGCCATGGTGGCGGCCCGGCAGCCCG   230

Qy    476 AGGAGATGAGCCTGCCCGGCAAGTGAAGGCCAAGATGATCGGCGGCGGCATCGGCGGCTTCA  535
Db    231 TGGTCAAGATCGAGGGTTCCACCCTCCGGTGCGGGTGGCCCGCTGCCTGCCTCCGCCA   290

Qy    536 TCAAGTGCGCCAGTACGACCAGATCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCA   595
Db    291 AGGACGCGGTCCCGGCGTGCCGGTCGAGCTGCAGAGGAGGGCCCGGCGTCAAGGCCA   350

Qy    596 CCGTGCTGATCGGCCCCACCCCGTGTGAACAATCATCGCCGCCGCAAACATGCTGACCCAGCTGG  655
Db    351 GCAGCGAGTGGATCCTCGACTGCATCGCCACGGCGGCGCATCTACGGCGTCACCACCG   410

Qy    656 GCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACGCGGCGCCGGCCCGTGAAGCTGAAGCCCG  715
Db    411 GCTTCGGCGGACCACTCCCACCCCGACCAAGGACGGCGGCCCGGCCCTCCAGGTCGAGCTGC   470

Qy    716 GCATGGACGGCGCCCAAGGTGAAGCTGAAGCAGTGGCCCCTGACCGGAGGAAGATCAAGGCCCTGA  775
Db    471 TCAGGCATCTCAACGCGGGAATCTTCGGCACGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACC  530

Qy    776 CCGCCATCTGCGAGGAGATGAGGGAGCATCATGGTGCGCGGCCGATGCTGGTGGCGCATGCTGGTGGTG  835
Db    531 AGGTTCACCAGGCGCCGCCATCAGCGGGCATCAACAACCCTCCTCAGGGTCTACTCCGGCA   590

Qy    836 CCTACACACACCCCGTGTTCGGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGG  895
Db    591 TCCGCTTCGAGATCCTGAGGCGATCACGAAGCTGCTCAACACCGGTGTCAACACCGGGTGTCAGCCCCTGCC  650

Qy    896 TGGACTTCCGGCGAGCTGAAGCAGCGCACCCAGGAGGTTCTGGAGGTTCAGCTGGCATCC  955
Db    651 TGCCGCTCCGGGCGCACCATCACCGGCTTGGGGCGACCTGGTCCCGCTCTCTACATCGCCG  710

Qy    956 CCCACCCCGCGCCGGCCTGCGAGGAAGAAGAGAGAGCTGACGTGGCGCGGCGACGCCT  1015
Db    711 GCCTCATCACGGCGCCGCCGGCCGTCACCG---TCGACGCGGGAGGAAGGTGG   767

Qy   1016 ACTTCAGCGTGCCCTGACGAGGACTTCCGCAAGTACACCGCGCTTCACCATCCCCAGCA  1075
```

  768 ACGCCGCGGAGGCGTTCAAGATCGCCGGCATCGCGGAGGCGGCGGTTCTTCTTCAAGCTCAACCCCA   827
Db
 1076 TCAACACGAGACCCCCGCATCGCATCCGTACCGTACCAACGTGCTGCCCAGGCTGGAAGG 1135
Qy
  828 AGGAGGGCCTCGCCATCGTCAACGGCACGTCCGTGGCTCCGCGGCTCGCGGCGCCACCGTGA   887
Db
 1136 GCAGCCCCAGCATCTTCAGAGCGAGCAGCATGACCAAGATCCTCGAGCCTTCGCGCCCCGCA 1195
Qy
  888 TGTACGACGCCAACGTCCTGGCCGTCCTGGTCCGGAGGTCCT---GTCCGCGGTCTTCTGCG   944
Db
 1196 ACCCCGAGATCGTGATCTACCAGGGCCCCCTGTACGTGGCAGCGGACCTGGAGATCGGCC 1255
Qy
  945 AGGTCATGAACGGCAAGCCGACTACACGGACACCTGACCCACAAGCTGAAGCACCACC 1004
Db
 1256 AGCACCGGCCCAAGATGAGGAGGGCCGGCGACCTGGAAGGGAGCTGGGGCTTCACCACCC 1315
Qy
 1005 CGGGTCATCGAGGCCGCATCATGGGAGGGCCATCCTGATGGGCGTACGAGCTGCACCCCCG 1375
Db
 1316 CCGACAAGGAGCACCAGAGAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCCG 1375
Qy
 1065 AGCAGGCCAAGAGTTGAACGAGCTGGACCCCGCTGCTGAAGCCCAAGCAGCAGGTACG 1124
Db
 1376 ACAAGTGAACCGTGCAGCCCATCGAGCTGCCCAGAAGGAGAGACCGTGAACGACA 1435
Qy
 1125 CGGTTCCCACGTCCGCCGCAGTGCGTGGGCTGGCCCCCAGATCGAGTC---ATCCGCGCCGCA 1181
Db
 1436 TCCAGAAGCTGGTGGCAAGCTGAACTGGGCCAGAGCCCCTTGGTGAGCATCACAAGGTGC 1495
Qy
 1182 CCAAGTCCATCGAGCGCGAGGTCAACTCCGTGAACGA---CAACCCGGTCATCGACGTCC 1238
Db
 1496 GCCAGCTGTGCAAGCTGCGTGCGGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCG 1555
Qy
 1239 ACCGGCGCAAGGGCGCTGCACGCGGCGCAACTTCAGGGCGTCAGGGGTGTCCATGG 1298
Db
 1556 AGGAGGCCGAGCTGGAGCTGTGAGCTGGCGGCCGAGAATCCGGCGAGATCGCGGAGCTGCACGGTCGACGTCTTCA 1615
Qy
 1299 ACACGCCCGCTTCGCCAACATCGGCAAGTCATGTTCGGCGAGTTCTCCGAGC 1358
Db
 1616 TGTACTACGACCCCAGCAAGGACACCTGGTCGGCCAAGAGCAGGCCACGACCAGT 1675
Qy
 1359 TCGTCAACGAGTTCTCACAACAACGGCTCACCTGACCCAACCTGGGCCGCGCAACCCCA 1418
Db
 1676 GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGATACGGCCAAGA 1735
Qy
 1419 GCCTGGACTTACGGCTTCAAGGGCACCATGAGCGCCTCCTACTGCTCCGAGCTCC 1478
Db
 1736 TGCGCACCGGCCACCAACGACGTGAGCGAGCAGCTGACCGTGCGAAGATCGCCA 1795
Qy
 1479 AGTACCTGGCAACCCCATCACCACCACGTGCAGAGCGCGAGCAGCACCAGGACG 1538
Db
 1796 TGGAGAGCAT---CGTGATCTGGGGCAAGACCCCCAAGTTCGGCCTGCCCATCCGAAGG 1852
Qy
 1539 TGAACTCCCTGGGCCTCGTCTCGGCGAGAAGACCGCGCAGAAGATCGGCATCGATCCTGAACC 1598
Db
 1853 AGACCTGGGAGAGACCTGGTGGACCTACTGCGCCGACGACTACTGGCCCCTGGACCTCGTGAGT 1912
Qy
 1599 TCATGTCGTCCACCTACATCGTGCGCTTGTGCCAGGCCGGTGCCACCTGCCAGCTGGCACCTCGGAGG 1658
Db
 1913 TCGTGAACACCCCCCCGTGGTGAACAGTCTGTGAAGTCGTCGTGAAGTGAGAAGGAGCCCCATCATCG 1972
Qy
 1659 AGAACATCAAGGCGTCGGTGAAGAACAACCGTGACCGACCAGGTCGGCCAAGAGGTGCTGACCA 1718
Db
 1973 GCGCCGAGACCTTCTACGTGGCACGGCGCGGCGCCGCCGAGATCGGCAAGGCCG 2032
Qy
 1719 TGAACCCCTCGGGCGGAGCTCTTCAGCGCCGAGAGATCGTGAGGCCTGACGAGACCAACCCCA 1778
Db
 2033 GCTACGTGACGACCGGGCCGGCGAGAGAATCGTGCTGTGACGTGGGCGCGACGCCT 2092
Qy
 1779 TCGACGGGAGCCGGTGTTCACGTACGGCGGCCGGCTCGACGCCTGCGCTGA 1838
Db
 2093 AGACCGAGCTGCACCGACGAGGACTTCCGCAAGTACACCGCCGCCCTTCACCATCCCCAGCA 2152
Qy
 1839 TGCAGAAGCTGCGGCCCGTGCTCGGTGGACCAGTACACACCGCGCCCTCAGCGCGGGCGACGCGGGAGCGCGG 1898
Db

```
Qy   2153  TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGCCCGGACACAGAGCGGAGAGCG 2212
                |||| |||||||||||||||||||||||||||||  ||||||||||  ||||
Db   1899  AGCCCTCCGTGTTCTCCAAGATCACCAGGTTCGAGGAGGAGCTCCGCGCGGTGCTGCCCC 1958

Qy   2213  AGCTGGTGAA 2222
                |||  |||||||
Db   1959  AGGAGGTGGA 1968
```

RESULT 2
BM321451
LOCUS   BM321451   951 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION   rockefeller.0.1211 Mastigamoeba balamuthi lambda ZAP II Library
              Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
              3.3.1.1), mRNA sequence.
ACCESSION   BM321451
VERSION   BM321451.1   GI:18055857
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM   Mastigamoeba balamuthi
              Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 951)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
              Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
              Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
              divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  MEDLINE   21819461
  PUBMED   11830664
COMMENT   Contact: Muller Miklos
              Laboratory of Biochemical Parasitology
              The Rockefeller University
              1230 York Avenue, New York, NY 10021, USA
              Email: mmuller@rockvax.rockefeller.edu
              Insert Length: 951   Std Error: 0.00
              POLYA=No.
FEATURES   Location/Qualifiers
     source   1..951
                  /organism="Mastigamoeba balamuthi"
                  /mol_type="mRNA"
                  /strain="ATCC 30984"
                  /db_xref="taxon:108607"
                  /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                  /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.5%;   Score 85.6; DB 12; Length 951;
Best Local Similarity   46.0%; Pred. No. 0.68;
Matches 323; Conservative   0; Mismatches 376; Indels   3; Gaps   1;

```
Qy   1728  CGCCAAGATGCGCACCGCCCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAA 1787
                 |||||||  |||   |||||||||
Db    219  CGCCAAGATCCGCTGGTGCTGCGTGCGCAACATCTTCTCGACGCAGGACCACGCGGCCGC 278

Qy   1788  GATCGCCATGGAGAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCA 1847
                 ||  |||||| |||||||||
Db    279  CATCGCGCAGGCCGCGGTCTCGGTCTTCGGTCTTCGGTGGAAGGGCCGAGAACCTCCAGGAGTACTG 338

Qy   1848  GAAGGAGACCTGGAGACCTGGTGGACCACTACTGGCAGGCCACCTGGATCCCGAGTG 1907
                 |||||||
Db    339  GGAGTGCACCTGGAAGGCCCTGTGCTTGCTTCCGGCCCTACCAGGCCCTACCAGGGCCCTACCGTCGA 398

Qy   1908  GGAGTTCGTGAACACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGAGAAGGAGCCCAT 1967
                 ||||||
Db    399  CGACGGCCGGTGACCGGACCTTGATGATCCACAAGGGGTTCGCCGGCCAGGACACCCCCAA 458

Qy   1968  CATCGGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCCAACCGGCAGGACCAAGATCGGCAA 2027
                 ||||||
Db    459  GCTGCTGGAGGACGACGAGGGCCTCGAGGAGGTCGCCTGCCTCCTCAACACGTGCTCAAGCA 518

Qy   2028  GGCCGGCTACGTGACCGACGACGGGGCGCCGAGAAGATCGTGAGCGTGACCGAGACCACCAA 2087
```

```
Db    519  GGTCCAGAAGGAGGAGCACGCCGGCTTCTGGCACAGATCCTCCCCGAGATCCGCGCGGTGTCAG 578
                |||||||||
Db   2088  CCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCAGCCAGCGGGCAGAGGTGAA 2147

Db    579  CGAGGAGACGACGACTGGCGTNATGAGGCTGTACCAGCTGCACCGCACGCGGCCGAAGCTGCT 638

Qy   2148  CATCGTGACCGACCAGCCAGTACGCTACCGCCCTGGGCATCATCCAGGGCCAGCCGGACAAGAGGCGA 2207
                |||||||||
Db    639  GTTCCCGGCCGTCAACGTCAACGAC---TCTNTCACCAAGAGCAAGTTCGAACACACATCTA 695

Qy   2208  GAGGCGAGTGGTGAACCAGCAACATCGAGCAGCTGATCAAGAACGAGAAGGTGTACCTGAG 2267
                |||||||||
Db    696  CGGCTGCCGCCACTCGCTCATCGACGGCATCAAGCGCGGCACCGACGTGATGCTCGGCGG 755

Qy   2268  CTGGGTGCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAA 2327
                |||||||||
Db    756  CAAGGTCGCGTCGTCGTCGCGGGGCTACGGCGACGTGGGCAAGGGCGTGCGCCGAGTCGCTGCG 815

Qy   2328  GGGCATCCGCAAGGTGCTGTTCCTGGACGGCGGCATCGATGGCGGCATCGTGATCATCACCAGTA 2387
                |||||||||
Db    816  CGGCCAGGGCTGCGCGTGCGCGTCATCGTGACGGAGATCGACCCCATCTGCGCGCTGCCAGGCGTC 875

Qy   2388  CATGACGACCTGTACGTGCGGCAGCGGCGCGCCCTTAGGATCGA 2429
                |||||||||
Db    876  GATGGCGCGGCTTCGAGGTCAACACGCTCGAGGCGGGGCTCGA 917
```

RESULT 3
CK159167/c
LOCUS   CK159167   869 bp   mRNA   linear   EST 05-DEC-2003
DEFINITION   FGAS040564 Triticum aestivum FGAS: TaLt5 Triticum aestivum cDNA,
              mRNA sequence.
ACCESSION   CK159167
VERSION   CK159167.1   GI:38985053
KEYWORDS   EST.
SOURCE   Triticum aestivum (bread wheat)
  ORGANISM   Triticum aestivum
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Pooideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 869)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
              Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
              Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
              Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE   Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Wm L Crosby
              Bioinformatics
              University of Saskatchewan, Department of Computer Science
              1C101 Engineering Building, 57 Campus Drive, Saskatoon,
              Saskatchewan, S7N 5A9, Canada
              Tel: 306 966 1769
              Fax: 306 966 2033
              Email: fgas_ests@cs.usask.ca
              This sequence is the direct result of the Base calling software
              Phred (default parameters). It is the raw base calls. To aid in the
              identification of the high quality insert the software Lucy
              (default parameters) has been run on this sequence. Lucy identified
              the region [128,636].
              Plate: TaLt537   row: N   column: 23.
FEATURES   Location/Qualifiers
     source   1..869
                  /organism="Triticum aestivum"
                  /mol_type="mRNA"
                  /cultivar="Wheat line PI 178383"
                  /db_xref="taxon:4565"
                  /lab_host="DH5 alpha"
                  /clone_lib="Triticum aestivum FGAS: TaLt5"
                  /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
                  subtractive hybridization) cDNA library from genotype
                  PI178383 cold hardened at 2 C for 21 days and 49 days
```

TITLE   Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
          Philippe,H.
          The analysis of 100 genes supports the grouping of three highly
          divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
PUBMED   11830664
COMMENT Contact: Muller Miklos
          Laboratory of Biochemical Parasitology
          The Rockefeller University
          1230 York Avenue, New York, NY 10021, USA
          Email: mmuller@rockvax.rockefeller.edu
          Insert Length: 1132   Std Error: 0.00
          POLYA=No.

FEATURES   Location/Qualifiers
   source   1..1132
               /organism="Mastigamoeba balamuthi"
               /mol_type="mRNA"
               /strain="ATCC 30984"
               /db_xref="taxon:108607"
               /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
               /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match   3.3%;   Score 80.2; DB 12; Length 1132;
Best Local Similarity 45.3%;   Pred. No. 3.1;
Matches 375; Conservative   0; Mismatches 444; Indels   9; Gaps   2;

Qy 139 CGGGCCCCGCAAGAAGGGCTCTGAAGTGCGGCAAGGAGGGCACCAGATGAAGGAC 198
         || ||| ||||||||||||| |||||||
Db   79 CGCCGCGGAGGGCAAGACGGACTACCGCGCGCGCCANCTGGTGATCCAGGACAAGAACAAG 138

Qy 199 TGCACCGAGGGCAGCCAACTTCTTCCGGAGGACCTGGCCTTCCCCAGGGCAAGGCC 258
                 |||||
Db 139 TACAACAGCCCCAAGTACCGCTTCGTC------GTCCGCTTCACCAACAGGGACATCGTC 192

Qy 259 CGGGAGTTCCCAGCGCGAGCGAGAACCGCGCCGGCCCCAACAGCCGCGAGCTGCAGGTG 318
         |||| |
Db 193 TGCCAGATCGCCTACGCCAAGATGACGGACGACCACATCTCGCCGCGCCTACTCGCAC 252

Qy 319 CGCGGCGACAACCCCGCTTCGGCCTCAAGCTCGGCCTGACCAACTACGGCGCCTGAACTTCCCC 378
          ||||||
Db 253 GAGCTCACCCGCTTCTGGCCGCCGTGTGCTCAGCTGCTGAAGAAGAGCTCAACCTGCCTACGGCGACT 312

Qy 379 CAGATCACCCTGTGGCAGCGCGCCCCCTGGTGAGCATCAAGGTGGGGCCAGATCAAGGAG 438
         ||||
Db 313 GGCCTGCTGTGCCGCCGTGTGCTAGAAGACACACCGTCTGAGGAGGAGAGTGAGCCTGCCGGCAAG 372

Qy 439 GCCCTGCTGACACACCGGCCGCATCGGCGGCATCGGCGCTTCATCAAGGTGCGCCAGTACGACCAG 498
         |||||
Db 373 GTCAAGAAGTCAACGGCGAGGACTACAACGTCGAGGAGCTCGACGAGGCTCGACGAGGCCCCCGT 432

Qy 499 TGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCGGCCGTGGTGTTCGCC 558
         |||||
Db 433 TTCAAGGCCCTGCTCGACGTCGGCCTGGTTCCGCACCTCACCGTGCTGCACCGTGATCGGCCCACCCCC 492

Qy 559 ATCCTGATCGAGATCTGGGCAAGAAGGCCATCGGCGGCGGGCTCAACGCTCCCC---CACAGCGAGACCCGCTTC 618
         |||||
Db 493 GCCCTCAAGGGCATGTGCCACGGCGGCGGCGGGCTCAACCGTCCCC---CACAGCGGAGACCCGCTTC 549

Qy 619 GTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGCTGCGCTGACACCTGAACTTCCCCATC 678
         |||||
Db 550 GTCGGCTTCACGGCCGACAAGAAGAGAGCTCAACGCCGCGGTCCTCCGCCAAGTACATCTTC 609

Qy 679 AGCCCCATCGAGACCATGCCCGTGCCGTGAAGCTGAAGCCCCAAGGTGAAG 738
         |||||
Db 610 GGCGGCCACGTCGCGCGTACATCAAGTCCTCCAAGGACGACCATCTGCGAGGAGATGGGAG 669

Qy 739 CAGTGCCCTGACCGAGGAGAGAATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGGAG 798
         |||||
Db 670 CGCCAGTTCTCGCCTACGCCAAGGAGAGGGTGTCACCGCCGACATGCTCGAGAAGATCTAC 729

Qy 799 AAGGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCC 858

Query Match   3.3%;   Score 80.6; DB 14; Length 869;
Best Local Similarity 44.9%;   Pred. No. 2.7;
Matches 305; Conservative   0; Mismatches 374; Indels   0; Gaps   0;

Qy 580 AAGAAGGCCATCGGCACCGTGCTGCTGATCGGCCCCACCCCCTGAACATCATCGGCCGCAAC 639
         |||||
Db 827 ATGGGCGGCGGCCCGCCAACCAGGATCACCACCAGGAGGACCAACCACCAACCACCACCAC 768

Qy 640 ATGCTGACCCAGCTGGCTGCACCCTGAACTTCCCCATCGAACTTCCCCATCGAGACCCATGCCC 699
         |||||
Db 767 CACCAACGACGACAACAGCGACACAACGACCACCAACACCAACAACAACAACAACAAC 708

Qy 700 GTGAAGTGAAGCCCGGCATGACGGCGCCCCAAGGTGAAGGCAGTGCCCCTGACCGAGGAG 759
         |||||
Db 707 ACCAACAACAACAACAACAACAACAACAACAACGACAACGACCACCACCACCAACAAAC 648

Qy 760 AAGATCAAGGCCCTGACCGCCATCTGCGAGGATGGAAGGAGGGCAAGATCACCAAG 819
         |||||
Db 647 AACAACAACGACAACAACAACAACAACAACAACAACAACAACAACAACAACAACACAC 588

Qy 820 ATCGGGCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCCAAGAAGAAGGACAGCACC 879
         |||||
Db 587 AACAACAACAACAACAACAACGACAACAACAACAACAACGACAACAACAACAACAAC 528

Qy 880 AAGTGGCGCAAGTCGTGACTTCCGCGAGCTGAACAAGGCGCACCCAGGACTTCTGGGAG 939
         |||||
Db 527 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACACAACAAC 468

Qy 940 GTGCAGCTGGGCATCCCCACCCCCACCCCCGGCGGCTGAAGGAGAGAGAGCGTGACCGTGCTG 999
         |||||
Db 467 AACAACAACAACAACAACAACGACGACAACGACGACGACGACAACGACGACGAC 408

Qy 1000 GACGTGGGCGACGGCCTACTTCAGCGTGCCCCTGGCCAAGTACACCCGCC 1059
         |||||
Db 407 GACAACAACGACAACAACGACAACAACAACGACAACAACAACAACAACAACAACAAC 348

Qy 1060 TTCACCATCCCCAGCATCAACAAGGACAGCGACATCTTCCAGCTACCAGTACAACGTGCTG 1119
         |||||
Db 347 GACCACAACGACAACGACAACAACAACAACAACAACAACAACGACAACAACCACAAC 288

Qy 1120 CCCGAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGAGCATGACCAAGATCCTGGAG 1179
         |||||
Db 287 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 228

Qy 1180 CCCTTCCGGCCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTTACGGTGGGCAGC 1239
         |||||
Db 227 AACAACAACAACAACGACGACAACGACGCAACGACGTCAACGATAACAACGACAACGACAAC 168

Qy 1240 GACCTGGAGATCGGCCAGC 1258
         |||||
Db 167 AACGGCAACAACAACAAAC 149

RESULT 4
BM320864
LOCUS   BM320864   1132 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.46 Mastigamoeba balamuthi lambda ZAP II Library
          Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
          sequence.
ACCESSION BM320864
VERSION   BM320864.1   GI:18005270
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM Mastigamoeba balamuthi
          Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE 1 (bases 1 to 1132)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,

Qy 390 GTGGCAGCGCCCCTGGTGAGCATCAAGGTGGGCGGGCCAGATCAAGAGGCCCTGCTGA 449
         ||||| |||||||||| | ||||| |||||| |||||||||||||||||||
Db 340 GGGCCGCCGTGTGCTGAAGAAGTCAACCTCTACTCCAGTAGTACGAGGGTGTCAAGAAGT 399

Qy 450 CACCGGCGCGACACACCGTGCTGGAGGAGGATGAGCCTGCCCGGCAAGTGGAAGCCCAA 509
         |||||||||| ||| ||||| |||||||| || |||||||| ||||||| ||||
Db 400 CAACGGCGGAGGACTACAACGTCGAGGAGCTCGACGACGGGCCCGCCGTTCAAGGCCC-- 457

Qy 510 GATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTAGCGCCGAGATCCTGATCGA 569
         |||| | |||| |||||||| |||||||||| |||||||| |||||||||||||||
Db 458 --TGCTGACGTCGGCGTCGGTCCGCACCTGACTGGCGCCCGTGTTCGCCGCGCCCTCAA 515

Qy 570 GATCTCGCGGCAAGAGAGGCCATCGCCGCACCGTGCTGATCGGCCACCCCGTGAACACATCAT 629
         |||| | ||||||||| |||||| |||| | || ||||||||||
Db 516 GGGCATGTGCGACGGCGGGCTCAAGTCCCCACAGCGAGACCCGCTTCGTCGGCTTCAA 575

Qy 630 CGGGCCCAACATGCTGACCCAGCTGGGCTGCACCTGACCTTCCCATCAGCCCCATCGA 689
         ||| || | ||| ||| ||| ||| |||||| |||||||||||||||
Db 576 CGGCGACAAGGAGGAGCTCAACGCGCGGTCCTCCGCAAGTACATCTTCCGGCGCCACGT 635

Qy 690 GACCGTGCCCGTGAAGCTGAAGCCCGGCCATCGGACGACGGCCCCAAGGTGAAGCAGTGGCCCCT 749
         ||||||| |||||||||| ||| ||| | ||||| |||||||||||||||||||| |||
Db 636 CGCGGCGTACATGAAGCTCCTCAAGGAGCAGGAGCCGCGCCCTTCGACCGCCAGTTCTC 695

Qy 750 GACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAA 809
         |||||||||||||||||| | | | || ||||||||| |||||||||| ||
Db 696 GCGCTACGCCAAGAGAGGGTGTCACCGCCGACATGTCGAGAAGATCTACACCGAGGCCCA 755

Qy 810 GATCACCAAGATCGGCCCCGAGAAACCCTACAAACACCCCGTGTTCGCCATCAAGAAGAA 869
         || ||| | || ||| || | |||||||||||||||||||||
Db 756 ------CAAGCAGATCCGCGACCGACCTTCGTCCCCAAGCGCCTCGAAGCCCGA 809

Qy 870 GGACAGCACCAAGTGCGCCAAGTCGTGTGGACTTCCGCGAGCTGACAGGCGACCACCCAGGA 929
         ||| | | ||| |||||| ||||||| | |||| |||| || ||||||||||||||
Db 810 GGGCGCCAAGCCCAAGCACTGGGCAAGCCGAGGCTGAGCGTACCACCAGAGGCCAAGAACCG 869

Qy 930 CTTCTGGAGGTGCAGCTGGCGCATCCCCACCCCCGGCTGAGAAGAAGAAGAGAGCGT 989
         ||||||||| |||||||||||||||| | || |||||||||| |||||||| |||||
Db 870 CGTCGCCCAGAGAAGAGTCCGCCTGGGCTACCCGACGCCCCAGAGCCAGTAAATTCC 929

Qy 990 GACCGTGTGGACGTGGGCGACGCCTACTTCAGCGTCGCCCTGACCGAGG 1039
         |||||| ||||| |||| ||| || | | ||| |||| ||| || ||
Db 930 CGGTGCCCTGTACACCCGCATCTCCCGCTGCGGCGGTTGCGCTGCTNGCGG 979

RESULT 6
BM321430
LOCUS   BM321430   867 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to ribosomal protein S4, mRNA
            sequence.
ACCESSION   BM321430
VERSION   BM321430.1   GI:18055836
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 867)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
   MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 867   Std Error: 0.00

Db 730 ACCGAGGCCCACAAGCAGATCCGCGACCGACCTTCGTCCCCAAGCGCGCCGCCTCGAAG 789

Qy 859 ATCAAGAAGAGGACAGCAGCACCAAGTGCGCCGCCAAGTCGGGCAGCTGGTGGACTTCCGCGAGCTGAACAAG 918

Db 790 CCCGAGGGGCGCGCAAGCCCAAGCACTGGGCAGCCGCAAGCGCGCAGGCTGCAGCGCTGACGTACCACGAGCGCCAAG 849

Qy 919 CGCACCCAGGACCTTCTCGGAGGTGCAGCTGGCGAGCGTGGGCGATCCCCACCCCGGC 966

Db 850 AACCGCGTCGCCCAGAGAAGAAGGTCCGCCTGGGCTACCCGGGCCCCACCCCGACGCCCCC 897

RESULT 5
BM320900
LOCUS   BM320900   1165 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
            sequence.
ACCESSION   BM320900
VERSION   BM320900.1   GI:18055306
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 1165)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
   MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 1165   Std Error: 0.00
            POLYA=No.
FEATURES   Location/Qualifiers
     source   1..1165
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.2%; Score 79.6; DB 12; Length 1165;
Best Local Similarity 45.1%; Pred. No. 3.7;
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;

Qy   90 CATCAAGTGCTTCAACTGCCGCAAGGAGGGCCACATCGCCCGCAACTGCGCGGCCCCCCG 149
         ||||||| |||||| | |||||||| || | | | |||||||||| ||||
Db   46 CGTCAAGAACAAGGGCTACTTCAAGGCGTTCAAGGCGTTCCAGACCCAGTTCCGTGCGCGGCCGCGAGGG 105

Qy 150 CAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACGATGAAGGAGACTGCACCGAGCG 209
         ||||||||||| ||| | ||||| |||||||||| ||||||| ||| |||| ||
Db 106 CAAGACGGACTACCGCGCGCGCCCCCCANCTGGTGATCCAGAACAAGCAAGTACAACAGCCC 165

Qy 210 CCAGGCCAACTTCTTCCGCGAGGACGTTCCCCCAGGCAAGCCCCGGAGTTCCC 269
         | | || | | | ||| ||||| |||||| ||||||||||
Db 166 CAAGTACCCGCTTCGTC-----GTCCGCTTCACCAACAGCAGGGACAATCGTCTGCCAGATCGC 219

Qy 270 CAGCGAGCAGCAGAACCGCGCCAACAGCCCACCAGCCGCGAGGTGCGCGGCGACAA 329
         ||| ||| | |||||||||||| | ||| | | ||| |||||
Db 220 CTACGCCAAGATCGACGCCGGCGGACCACACATCCTGCGCGCGCCTACTCGCAGAGCTCACCCG 279

Qy 330 CCCCCGCAGCGAGGGCCGGCGCCGGGCCGCCGAGCGCGCAGGGCACCCGGACCACCCT 389
         |||||| ||| || |||||| || || | | ||| |||| |||||| |||||||
Db 280 CTTCGGCGTCAAGCTCGGCCTGACCAACTACGCGCCGCGCCTACGCGGCGACTGGCCTTGCTGCT 339

SOURCE   Mastigamoeba balamuthi
  ORGANISM Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 1550)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 1550   Std Error: 0.00
            POLYA=No.

FEATURES   Location/Qualifiers
     source   1..1550
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match   3.1%; Score 76.6; DB 12; Length 1550;
Best Local Similarity 45.6%; Pred. No. 8.9;
Matches 308; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

Qy   1755 CGAGCTGAAGCAGCAGCCGTGACCGAGGCCGTCAGAAGATCGGCATGGAGAAGATCGTGATCTG 1814
          ||| | || | || | || || | || |||||||| |||||| | ||||||||||| | |
Db   5 CGGCACGAGGACGCAGGACCACGCGGCGCCGCCGCATCGGCATGGGAGAAGATCGTGATCTT 64

Qy   1815 GGGCAAGACCCCCAAGTTCCGCCTGCCATCCAGAAGGAGACCTGGGAGACCTGTGGAC 1874
          | || | | ||||| | | ||| |||| |||
Db   65 CGCCTGAAGGGCGGAAACCTTCCAGAGTACTGGGAGTGCACCTGGAAGGCCCGTGCTT 124

Qy   1875 CGACTACTGGCAGGCCCTGGCACCTGATCCCGAGTGGGAGTTCGTGACGACGGCGGTGGT 1934
          | | | | || || || || | | || | | || | || || | | | |||
Db   125 CGGCCCCTACCAGGGCCCTACAACCAAGTCGACGATCATCGTCGACACGGCGGTCTGATGAT 184

Qy   1935 GAAGCTGTGGTACCAGCTGGAGAAGGACCCATCATCGGCGCCGAGACCTTCTACGTGGA 1994
          || | | | | || | || || | | || | ||| | | | || |
Db   185 CCACAAGGGGTTCGGCGGCAACCACCCAAGCTGCTGGAGGACGACGAGGGGGCCTCGA 244

Qy   1995 CGGCGCCGCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGCCG 2054
          ||| | | || ||| | | | | | || | || | || | || ||
Db   245 GGAGGTGCCTGCCTGCCTCAACACGTGCTTCAACGCAGGTCCAGAAGGAGCAGCCCGGCTTCTG 304

Qy   2055 GCAGAAGATCGTGAGCCTGACCGAGACCAAACCAGAAGACCGAGCTGCAGGCCATCCA 2114
          ||| || | | || | | ||| || || | || | || || | ||| | |
Db   305 GCACAAGATCCTCCCCGAGATCGGCGGTGTCAGCGAGGAGATGACGACTGCCGTGATGAG 364

Qy   2115 GCTGGCCCTGCAGGACAGCCCGTGACCGAGGTGAACATCGTGACGGCCCT 2174
          || | || | | | | | | | | || | | | | | || | |
Db   365 GCTGTACCAGTCGCACCGGACGCAAGCTGCTGTTCCCGGCGTCAACGACTC 424

Qy   2175 GGGCATCATCCAGGCCGAGCTGGTGGTGAACCAGATCATCGA 2234
          | | || | || | | | | | | || || | |||
Db   425 TG---TCACCAAGAGCAAGTTTGACAACATTCACGGCTCATCGCAGGG 481

Qy   2235 GCAGCTGATCAAGAAGGAGAAGTGTACCTGAGCTGGTGGTCCGCCCACAAGGGCATCGG 2294
          | | | || || ||| | | || ||| | | | ||| ||| | | ||
Db   482 CATCAAGCGCGGCACCGACGTGATGCTCGGGCCGCAAGGTCGTCGCGGTCGCGGGCTACGG 541

Qy   2295 CGGCAAGCAGCAGCAGATCGACAAGCTGGTGACGAAGTGCTGTTCCTGGA 2354
          | | | | | | | | | | || | | | || ||||
Db   542 CGACGTGGGCAAGGCGCTGCGCCAGTCCTGCTCGCCGGGTCGGGGTCATCGTGAC 601

FEATURES   Location/Qualifiers
     source   POLYA=No.
                     1..867
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"

ORIGIN

Query Match   3.2%; Score 79; DB 12; Length 867;
Best Local Similarity 45.7%; Pred. No. 4.1;
Matches 314; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

Qy   58 CAGCGCAGCAACTTCAAGGGCCCCAAGCGCGATCATCAAGTGCTTCAACTGCGGCAAGGAG 117
          || | | | | ||| | ||| | | | |||| | || | || | ||||||||| |
Db   131 CCGCACAAGATGCGGCGAGTGCTGCTGCCGTCAATCATCCTGCTGCGCAACAGGTTGAAGTAC 190

Qy   118 GGCCAATCGCCGCAACTGCCGCGCCGCCCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAG 177
          |||| ||| | | |||| | || | | | | | ||| | | || |||| | ||
Db   191 GCGCTGACCCGCCGTGAGGTCACCTGCAACGTGATCGAGCCCTGATCAAGATCGCACGGC 250

Qy   178 GAGGGCCACCAGATGAAGGACTGCACCGAGGCGCAGGCGCCAACTTCTTCCGGGAGGACCTG 237
          | | || || | ||| | | | | | | || || | ||| || || |
Db   251 AAGGTCCGCACCGACACGCCACTTCCCCGCGGCTTCATGGACGTGCTCTCGATCGACAAG 310

Qy   238 GCCTTCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCC 297
          || | | || | |||| | ||| ||| || |||| | | | || ||| || ||
Db   311 ACGGACGAGCACTTCCGCGCTGCTGTCTACGACACCAACGGGCCGCTTCCAGGCGCCACCGCATC 370

Qy   298 ACCAGCCCGGAGCTGCAGGT---GCGCGGCGACAACCCCGCAGCGAGGCCCGGCGCGAG 354
          | || | || | ||| | |||| | | | ||| || || || | |
Db   371 AACTCGACCGAGCCCAAGTTCAAGGTCCGCGCGTGCAGTCCGGCAACAAG 430

Qy   355 CGCCAGGGGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATC 414
          || | || | | || ||| || | | | ||| | | || | | | | || |||| ||
Db   431 GGCATCCGTACCGGTGACCCACGAGACGCGCGAGATCGCGTACCCCGACACCCCAACCCGACATC 490

Qy   415 AAGGTGGGCGGGCCAGATCAAGGAGGGCCCTGCTCGGACACCGGCCGTGCTG 474
          || | | | || | | | |||| || | | | || | | |||
Db   491 AAGGTCAACGACACGGTCAAGATCGACCTCGCCTGGCAACAGATCATCGTCGACTTCGTCAGG 550

Qy   475 GAGAGATGAGCCTGCCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCGTTC 534
          || | || ||| || | || | | | || | |||| | | | |
Db   551 TTCGGATCGGCAACCTCGTCATGATCACTGCGGGCCAACCTTGGCCGCGTCGGGGTC 610

Qy   535 ATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGCCAAGAAGGCCATCGGC 594
          | | | || | | || ||| | | | || | | || || | | || | |
Db   611 ATTGTGCGCGGCGAGAAGCACGCGAGGGCTCGTTCCAGATCATCCACGTCAAGGACGCGCCGTC 670

Qy   595 ACCGTGCTGATCGGCCCCACCCCCGTGAACATCACGCCGCAACATCGTGACCCAGCTG 654
          ||| | | | | | | | ||| || | | || | ||| | || || | |
Db   671 GGCCACCAGTTCGGCGACGCGGCCTGACGCGGCCTGACCAAGGGCACCAAGTCC 730

Qy   655 GGCTGCACCCGAACTTCCCCATCAGCCCCATCGGACAAGGGCATCAAGAGTCAAGAGTTCCAGGCG 714
          ||| | | || | | | ||| || || || || | || | || | | || ||
Db   731 CTCGTCACGGTCGCCCCGGCCCGGCGGCCATCAAGAGTCGATCATCATGAGGAGTTCCAGGCG 790

Qy   715 GGCATGGGACGGCCCCAAGGTGAAGCAG 741
          |||| || | | || | ||| ||| |
Db   791 CGCCACGCGGCCACGAGGACCCAGGAGGAG 817

RESULT 7
LOCUS   BM321022   1550 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION   BM321022
VERSION   BM321022.1 GI:18055428
KEYWORDS   EST.

```
RESULT 8
BM321023
LOCUS       BM321023              862 bp    mRNA    linear   EST 03-JAN-2002
DEFINITION  rockefeller.0.594 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION   BM321023
VERSION     BM321023.1  GI:18055429
KEYWORDS    EST.
SOURCE      Mastigamoeba balamuthi
  ORGANISM  Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1  (bases 1 to 862)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., and
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE     The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  MEDLINE   21819461
   PUBMED   11830664
COMMENT     Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 862  Std Error: 0.00
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1..862
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match          3.0%;   Score 75;  DB 12;  Length 862;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 283;  Conservative 0;  Mismatches 327;  Indels 3;  Gaps 1;

Qy 1728 CGCCAAGATGCGCACCGCCACCGCCCACCAACGACGTGAAGCAGCTGACCGGACGGCCGTGCAGAA 1787
Db  220 CGCCAAGATCCGCTGGTGTCCGTCGCGCTCGCAACATCTTCTCGCACGCAGGAGCCACGCGGCGCCGCC 279
Qy 1788 GATCGGCCATGGAGGAGCATCGTGATCGTGGGCAAGACCACCCCAAGTTCCGCCTGCCCATCCA 1847
Db  280 CATCGCGCAGGCCGGAGTCTCGGGTCTTCGGTCTTCCGCCTGGAACGGGCGAGAACCTTCCAGGAGTACTG 339
Qy 1848 GAAGGAGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCCAGGCGACTACTGGCCAGGCGACTCCCGAGTG 1907
Db  340 GGAGTGCACCTNGAAGGGCCCTGTGCTTCGGCCCTACCAGGGACCTGGAGGAGGAGCCCAT 399
Qy 1908 GGAGTTCGTGAACACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGGAGGAGCCCAT 1967
Db  400 CGACGGCGGTGACGGCGACTCTAATGATCCACAAGGGGTTCGGCGGCCAGGACCAACCCCAA 459
Qy 1968 CATCGGCGCCGAGACCTTCTACGTGGACGGCGGCGCCGCCAACCGGCAGACCAAGATCGGCAA 2027
Db  460 GCTCCTGAGGACGACGAGGGGCCTCGAGGAGGTCGCCTTCGGCCTTCAACAACGTGCTCAAGCA 519
Qy 2028 GGCCGGCTACGTGACCGACGACGGGGGCGGCCGGCGCGAGAAGATCGGCAGAAGATCGGCACCACCAA 2087
```

```
Qy  520 GGTCCAGAAGGAGCAGCCCGGCTTCTCGCACAAGATCCTCCCGAGATCCGCGGTGTCAG 579
Db 2088 CCGAGAAGACCGAGCTGCAGGCCATCCAGCTGCCCTGCCAGGACAGCGGCAGCGAGGTGAA 2147
Db  580 CGAGGAGACGACGACTGGCGTGATGAGCGTGACCCTGGGCATC--ATCCAGGCGCCAGCCCAGCAAGCTGCT 639
Qy 2148 CATCGTGACCGACAGCCAGTAGCGCCCTGGGCATC--ATCCAGGCGCCAGCCCAGACAAAGAG 2204
Db  640 GTTCCCCGGCCGTCAACGTCAACGACTCGTCGTCACCAAGAGCAGCAAGTTTGACAACATCTACGG 699
Qy 2205 CGAGAGCGGAGCTGGTGAACCAGATCATCGAGCAGTCGATCAAGAAGAGAAGGTGTACCT 2264
Db  700 CTGCCGCCACTCGCTCATCGACGGCATCAAGCGGCCGCATCAAGCGACCGTGATGCTCGGCGGCAA 759
Qy 2265 GAGTCGGTGGCCCCCCCACAAGGGCATCGGCGGCAACGAGCAGCAGATCGACAAGCTGGTGAG 2324
Db  760 GGTCGCCGTCGTCGCGGCGTACGGCGACGTGGGCAAGGGCTCGCGCCGCCAGTCGTCGTCGCGGG 819
Qy 2325 CAAGGGCATCCGC 2337
Db  820 CCAGGGCTGCCGC 832

RESULT 9
BZ568300
LOCUS       BZ568300                914 bp    DNA    linear   GSS 17-DEC-2002
DEFINITION  pacs2-164_7463.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_7463, genomic survey sequence.
ACCESSION   BZ568300
VERSION     BZ568300.1  GI:27201058
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 914)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Psedomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..914
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_7463"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN

Query Match          3.0%;   Score 74.4;  DB 28;  Length 914;
Best Local Similarity 47.0%; Pred. No. 15;
Matches 265;  Conservative 0;  Mismatches 296;  Indels 3;  Gaps 1;

Qy 652 CTGGGCTGCACCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGTGAAG 711
Db 112 CTGCGCAACCAGCAGCAGCGGCCTGATCTTCAGCCAGGCCGCCCTGACGGTGCAGGCC 171
Qy 712 CCCGGCATGGACGGCCCCAAGCGTGAAGCAGTGGCCCCTGACGGAGGAGATCAAGGCC 771
Db 172 GGCAGCCTGACCAACCCGCAACCGCGACCGCAGCGGCGACAACCAGCGTGCCGTATC 231
```

```
Matches   299;   Conservative   0;   Mismatches   329;   Indels   9;   Gaps   2;

Qy   388  CTGTGGCAGCGCCCCTGGTGAGCATCAAGTGGGGCGCCAGATCAAGGAGGAGGGCCCTGCTG  447
               ||||   ||||      |||||
Db   125  CTGGCCGCCGTGTGTCGAGAAGCTCAACCTGACTCCAAGTACAGAGGGTGTCAAGAAG  184

Qy   448  GACACCGGCGCCACGACACCCGTGCTGGAGGAGATCAGCCTGCCCGGCAAGTGGAAGCCC  507
               |||   |||||||
Db   185  GTCAACGGCGAGGACTACAACCTGAGACTCGAGGAGGAGTCCCCGTTCAAGGCC  244

Qy   508  AAGATGATCGGCGACGGCTTCATCAAGGTGCGCAAGATCCTGATC  567
               ||||||
Db   245  CTGCTCGACGTGCCGGCCTGGTCCGCACCTGACCTGCCGCGGCGCCCGGTGTTCGCCGCCTCAAG  304

Qy   568  GAGATCTGCGGCAAGAGAGCCATCGGCAAGAGAGGCGTGCACCGGCGCCCCACCCCGTGAACATC  627
               ||||
Db   305  GGCATGTGCGACGGCGGCGTCAACGTCCCC---CACAGCGAGACCGCGCTTCGTCGGCTTC  361

Qy   628  ATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGACCTTCCCCATCAGCCCCATC  687
               ||   |||||||||
Db   362  AACGGCGACAAGAAGGAGCTCAACGCCGGGTCCTCCGCAAGTACATCTTCCGGCGGCCAC  421

Qy   688  GAGACCGTGCCCGTGAAGCTGACAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCC  747
               ||   ||||||
Db   422  GTCGCCGCGTACATGAAGCTCCTCAAGGAGCAGGAGCGCCGCCTTCGACCGCCAGTTC  481

Qy   748  CTGACCGAGGAGAGATCAAGGCCCTGACCGCCCATCTGGAGGAGTGGAGAAGAGGGGC  807
               |||||||
Db   482  TCGCGCTACGCCAAGGAGGGTGTCACCGGCGACCAGTCTGGAGAAGATCTACACCGGAGCC  541

Qy   808  AAGATCACCAAAGATCGGCCCCGAGAACCCTACAAACACCCCGTGTTCGCCATCAAGAAG  867
               ||||||||
Db   542  CA------CAAGCAGATCCGGCGACGCAAGTCGTGTGGACTTCCGGAGCTGAACAAGCGCACCCAG  595

Qy   868  AAGGACAGCACCAAGTGCGCAAGTGCTGTGGCAGCTTCCGGCTGAGCTGAACAAGCGCAGC  927
               |||||||
Db   596  GAGGGCGCCAAGCCCAAGCACTCGGGGCAAGGCTGACGTACCTACCACGAGCGCAAGAAC  655

Qy   928  GACTTCTGGAGTGCAGCTGGCATCCCCACCCCGCCGGCTGAAGAAGAAGAAGAGAGC  987
               |||||||
Db   656  CGCGTCGCCCAAGGAAGGTCCGCTCGGCGTACCCCGAGCGCCCCAAGAGCGCCCAGTAATT  715

Qy   988  GTGACCGGTGCGACGTGGGGGACGCGCTACTTCAGCG  1024
               |||||||
Db   716  CCCGGTGCCCTGTACACNGCCATCTCCCGCTGCGGCG  752
```

RESULT 11
BM587428/c
LOCUS   BM587428   566 bp   mRNA   linear   EST 25-FEB-2002
DEFINITION   17000687321202 A.Gam.ad.cDNA.bloodl Anopheles gambiae cDNA clone
             1960044969610 5', mRNA sequence.
ACCESSION   BM587428
VERSION   BM587428.1 GI:18883289
KEYWORDS   EST.
SOURCE   Anopheles gambiae (African malaria mosquito)
ORGANISM   Anopheles gambiae
             Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
             Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
             Anopheles.
REFERENCE   1 (bases 1 to 566)
AUTHORS   Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L.,
             Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE   Celera Anopheles gambiae EST project
JOURNAL   Unpublished (2002)
COMMENT   Contact: Holt R.A.
             Celera Genomics
             45 w. Gude Dr., Rockville, MD 20850, USA
             Tel: 2404533151
             Fax: 2404534580
             Email: HoltRA@celera.com
             Plate: NU01004AAX row: G   column: 08
             Seq primer: M13 Reverse.

```
Qy   772  CTGACCGCCATCTCGGAGGAGGATGGAGGAGGAGGGCCAAGATCACCAAGATCGGCCCCGAG  831
               ||||      ||||
Db   232  GGTGGCGCCTGGACAACACCCCGTGTTCGCCATCAAGAAGGCCGGT AACCTAGACCTG  291

Qy   832  AACCCTACAACACCCCGTGTTCGCCATCAAGAAGGAGGACACAGCACCAAGTGGCGCAAG  891
               |||      ||||||
Db   292  CAGAGCGGCAGCCTCGACAACGGCCTGCCTGCTCAACAGCGCCAAGGGTTGGCTG  351

Qy   892  CTGGTGGACTTCCGGCAGTGAACAAGCGCACCCAGGACTTCTCGGAGGTGCAGTCGGGC  951
               ||      |||||||
Db   352  AAGCTGGTCACCGGGCTGTTCGACAACGAGCGGTCACCCAGGCGGCAGTCGCTGAA  411

Qy   952  ATCCCCACCCCGCCGGCGCCTGAAGAAGAGAGCGTGACCGGTGCTGACGTGGGCGAC  1011
               ||
Db   412  ATCCGGCGGGCAAGAGGCGTGCGCAAGGGCGTGCGCAACCAGCAGAGGC---CATGTCGGGCGCGCGGGGGGC  468

Qy   1012 GCCTACTTCAGCGTGGCCCCTGGACGAGGACTTCCGCAAGTACACCCGCCTTCACCATCCCC  1071
               ||      |||||||
Db   469  GACAACCGCATCGTCACCGTCGACTTCGACAACCAGGTGGCGGCCTCTACGCCAGCGGC  528

Qy   1072 AGCATCAACAACGAGACCCCGGCCATCCGCTACCAGTACACAACGTGCTGCCCCAGGGCTGG  1131
               ||      ||||
Db   529  CTGCTCAGCCTCGACGGCGCAGCGCTTCCTCAACCAGGGGCGCGGCGGCCAGGCGGCGGC  588

Qy   1132 AAGGGCAGCCCAGCATCTTCCAGAGCAGCAGCAAGATGACCAAGATCTTCAGCCTTCCGCGCC  1191
               |||||      ||||
Db   589  AAGGTCGGCGACGGCGCAGCGCATCGACTTCAGCCTGGCCGGCCGGCCTGGCCAACCCGCTTCGCC  648

Qy   1192 CGCAACCCCCGAGATCGTGATCTAC  1215
Db   649  AGTTGGAAAGCGAGAGCGAGCTGC  672
```

RESULT 10
BM321393
LOCUS   BM321393   853 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION   rockefeller.0.1222 Mastigamoeba balamuthi lambda ZAP II Library
             Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
             sequence.
ACCESSION   BM321393
VERSION   BM321393.1 GI:18055799
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
ORGANISM   Mastigamoeba balamuthi
             Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 853)
AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., and
             Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
             Philippe,H.
TITLE   The analysis of 100 genes supports the grouping of three highly
             divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE   21819461
PUBMED   11830664
COMMENT   Contact: Muller Miklos
             Laboratory of Biochemical Parasitology
             The Rockefeller University
             1230 York Avenue, New York, NY 10021, USA
             Email: mmuller@rockvax.rockefeller.edu
             Insert Length: 853   Std Error: 0.00
             POLYA=No.
FEATURES   Location/Qualifiers
     source   1..853
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.0%;   Score 73.2;   DB 12;   Length 853;
Best Local Similarity 46.9%;   Pred. No. 20;

University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 03 row: L column: 13
Seq primer: gta aaa cga cgg cca gtg.

```
FEATURES             Location/Qualifiers
     source          1..788
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEb03L13"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEb"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (Che
                     86061)"
ORIGIN
```

Query Match   2.9%;   Score 72.6; DB 14;   Length 788;
Best Local Similarity   46.3%;   Pred. No. 24;
Matches 314; Conservative   0; Mismatches 355; Indels 9; Gaps 2;

```
Qy    107 GCGGCAAGGAGGGGCCACATCGCCGCCAACTGCCGCGGCCCCCGCAAGAAGGGCTGCTGGA 166
              ||||   |  || |||||    ||   ||| ||  || | | |  ||||||| |||||
Db     99 GGCGGCGGCGGGCCGCGGAGATGGCCGGCGCCGCCAGCCCACCTGCAGCGTGAGGCGTGCGC 158

Qy    167 AGTGCGGCAAGGAGGGCCACCAGATGAAGACTGCACCGAGCGACCCGATGGAGCAGAAGGG 226
              || ||||  | || || | || ||| || ||   |   | ||   | || ||| ||
Db    159 AGTTCCGCGAGCCGCTGGTCAAGATCCAGGCGGCGCCACCCTCCGCGTCCGGCGGTGGCCG 218

Qy    227 GCGGAGGACCTGGCTTCCCCAGGGCGCAAGGCCCGCCGAGTTCCCAGGCAGCAGAACCGCG 286
              |||||  |||||| |      |||||  || ||    | |||  | |  ||| | |
Db    219 CCGTCGCCCAGGCCAAGGACCGCGTGTCGTCGTCGTCGCCGCGCAGCGAGGGCCGCC 278

Qy    287 CCAACAGCCCCCACCAGCCGCAGCTGCAGGTGCGCGGCGCAACCCCGCAGCGAGGCCG 346
              | | | ||| |||| |  ||| |||  |||| | |||| ||||   |||| | || |
Db    279 CCGGTCAAGGCCGCAGCGAGTGGATCCTCAACTGCATGCCCCACGGCGCGGCGACATCT 338

Qy    347 GGGCCGAGCGCCAGGGCACCTGAACTTCCCCGAGATCACCCTGTGCGACCAGCCCCTGG 406
              ||| |||| | ||| ||| | | | | | ||  | |  || |||  | | | ||
Db    339 ACGGCGTCACCACCGGCTTCGGCGGCCACCTCCACCGCGCACGGACCGGCCCCGCCC 398

Qy    407 TGAGCATCAAGGTGGGGCGGCCAGATCAAGGAGGGCCCTGCTGACCACCGGCGCCGA--CG 463
              | |  |||||   |||   | || |||     ||| |  |   |||| |  ||   |  ||
Db    399 TCCAAGTCGAGCTCCTCAGGCATCTCAACGCGCGGCCGGAATCTTCGGCCACTGGCTCCGATGGCC 458

Qy    464 ACACCGGTGCTGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCAAGATGATCGGCGGCA 523
              | || ||| ||   | || || || ||  |||||| ||  | || | ||| ||| ||
Db    459 ACACGCTCCGTCGGAGACGGTGCGGCGGGCCATGCTCGTGCGGCATCGAACACCCTCCTCC 518

Qy    524 TCCGGCGGCTTCATCAAGGTGACCGCCAGTACGACCAGATCCTGATCGAGATCTGGCGCAAGA 583
              |  ||||| || |||| | | |  ||  |   ||| |  |  ||  ||   || || |
Db    519 AGGGCTACTCCGGTTCGAGATCCGGTTCGAGATCCGGCCATCACCGAGGCCGCATCACCGTGCTCAACACCG 578

Qy    584 AGGCCATCGGCACCGTGCTGATCGGCCCACCCGTGAACATCATGGGCCGCAACATGC 643
              | ||||  | || ||    || |||||    ||| |  || || |||||| ||   |
Db    579 GCGTCAACGGTCGCTCCGGCTCCGTGCACCATTCACCGGCGTGCGGTGACCTGGTTCCCC 638

Qy    644 TGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACGGTGCCCGTGA 703
              ||    || |||| || |    || | | ||| |  |||||   |  |||   ||   |||
Db    639 TGTCCTACATTGCCGGCCTCATCACCGCCCCCAAGCGGGCCATCTCGCCCCACG 698

Qy    704 AGTGAAGCCGGCCATGACGGCCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGA 763
              ||||    ||  |   ||| |||| ||||    |   ||  || |  ||  |||| ||
```

```
FEATURES             Location/Qualifiers
     source          1..566
                     /organism="Anopheles gambiae"
                     /mol_type="mRNA"
                     /strain="RSP-ST (Reduced susc. to Permethrin - std.
                     chromosome)"
                     /db_xref="taxon:7165"
                     /clone="196004496110"
                     /dev_stage="Adult"
                     /lab_host="DH10b"
                     /clone_lib="A.Gam.ad.cDNA.blood1"
                     /note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole
                     adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
                     hours after human blood feeding. cDNA inserts >500 bp
                     cloned directionally into pSport 1. Not 1 site is 3'.
                     Clones available through the Malaria Research and
                     Reference Reagent Resource Center (www.malaria.mr4.org)"
ORIGIN
```

Query Match   3.0%;   Score 72.8; DB 12;   Length 566;
Best Local Similarity   48.3%;   Pred. No. 21;
Matches 203; Conservative   0; Mismatches 217; Indels 0; Gaps 0;

```
Qy   1950 GCTGGAGAAGGAGGAGCCCATCATCCGCGCCGAGACCTTCTACGTGGACGGCGCGCCAACCG 2009
              | | ||| | ||| |  ||||   || ||||  || ||  |  | ||| || |   | ||
Db    482 GTTCGAGATCGACGCCACCAACGGCATCCTGCAGGTGTCGGCCAGGTGCAAGGGCACGGGCAA 423

Qy   2010 CGAGACCAAGATCGGCAAGGCCAAGGCCCGCTACGTGACCGACCGGGGCGGCGCAGAGATCGTGAG 2069
              | || || ||| |||  ||| | || ||| || ||||  | ||  |||  | | ||  | |
Db    422 CCGGGAAAAGATCGTCATCCACCAACCAGCCGAGAACCGGCCTGACCGCCGACGCATCGAGCG 363

Qy   2070 CCTGACCGGACCACCACCAGAAGACCCAGCTGCAGGCCATCCAGTTGCCCTGCAGGA 2129
              | ||| | || || ||||| || |||| || ||   |  |  | ||| ||| |||
Db    362 CATGATCAAGGATCGGACGCGGTTCGCCGACACGCAAGAGCTGAAGAGGAGCGCGTGGA 303

Qy   2130 CAGCGGCAGCGAGGTGAACATCGTGACGAGTGAACATCGATACGCCCTGGGCCATCATCCAGGC 2189
              |||||  |  || |  || | || ||||| || | |   ||  | ||  || | ||| ||||
Db    302 GGCCCGCAACGAGCTCGAGACTTCGAGAGCTACGGTACAGCTCAGCGTGAGAACCAGCTCAGCTCGAAGGA 243

Qy   2190 CCAGGGCCCGACAAGAGGCGAGAGAGCGAGCTCGTGGTGAACCAGAGTCATCATGGAACCAGATCAAGAA 2249
              |||||  | |   |   | |  |  | || | | || |  ||  || | | | |  || |   |
Db    242 CAAGCTGGGCGGCGAGCGTGTCCGACGACGACAAGGCCCAAGATGGAGGGCGATGACGA 183

Qy   2250 GGAGAAGGTGTACCTGAGCTGGGTGCCCCCACCAAAGGGCATCGCGGCAACGAGCAGCAGAT 2309
              |||||    |  |  ||| ||| || | |||  ||  | |||| |  | ||| || |
Db    182 GAAGATCAAGTGGCTGGACGAGAACCAGGACACCAGGCCGGAGAGTACAAAGAAGCAGAA 123

Qy   2310 CGACAAGCTGGTGAGCAAGGGCATCGCGCAAGGTGCTGTTCCTGACCGGCATCGATGGCGG 2369
              | | |||| || | || ||| ||| | | |  || || |  | ||| | ||  |||  ||
Db    122 GAAGGAGCTGGAGGACATCTGCAGCCCATCATTGCCAAGTCGTACGCGGCAGTGGGCG 63
```

```
RESULT 12
CB643171
LOCUS       CB643171                 788 bp    mRNA    linear    EST 08-APR-2003
DEFINITION  OSJNEb03L13.f OSJNEb Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEb03L13 5', mRNA sequence.
ACCESSION   CB643171
VERSION     CB643171.1  GI:29638162
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 788)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
  COMMENT   Contact: Rod Wing
            Arizona Genomics Institute
```

Db   699 GCAGGAAG------GTGGACGGCCGCCGAGGCGTTNCAGCTCGCCGGCATCGAGGGTGGCT 752

Qy   764 TCAAGGCCTGACCGCCA 781
             ||| |||||| |||
Db   753 TCTTCACGTGAACCCCA 770

RESULT 13
CC675888
LOCUS   CC675888   753 bp   DNA   linear   GSS 19-JUN-2003
DEFINITION   OGWCO51TH ZM_0.7.1.5 KB Zea mays genomic clone ZMBMA0539J05,
             genomic survey sequence.
ACCESSION   CC675888
VERSION   CC675888.1 GI:32080584
KEYWORDS   GSS.
SOURCE   Zea mays
  ORGANISM   Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 753)
  AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
             Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
             Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  TITLE   Consortium for Maize Genomics
  JOURNAL   Unpublished (2002)
COMMENT   Other GSSs: OGWCO51TV
             Contact: Cathy Whitelaw
             TIGR
             9712 Medical Center Drive, Rockville, MD 20850, USA
             Tel: 301-838-5843
             Fax: 301-838-0208
             Email: whitelaw@tigr.org
             Seq primer: TR
             Class: sheared ends.
FEATURES   Location/Qualifiers
     source   1..753
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMBMA0539J05"
                     /clone_lib="ZM_0.7.1.5_KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"
ORIGIN

Query Match   2.9%; Score 72.4; DB 29; Length 753;
Best Local Similarity 46.2%; Pred. No. 25;
Matches 281; Conservative 0; Mismatches 321; Indels 6; Gaps 1;

Qy   31 CAGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCAAGGGCCCAAGCGCCATC 90
            ||| || || ||||| ||| | | || || || || |||| |||| |||||
Db   40 CACACCACCGACGACACCTTCCGCGGGGTCCGCGTCAAGTGGACCTCCACCACGCGCACC 99

Qy   91 ATCAAGTGCTTCAACTGCCGCAAGGAGGGCGCGGGCGGGGTAACCGTACAACATATTCGG 150
            | | ||||||||| || ||| || || || || | || || ||| ||| |
Db   100 GTCGACCGGCAGCGGCAGCGGCGGCGCGGCGGGTAACCGTAAGCAAGGCCGGCGGCGGC 159

Qy   151 AAGAAGGGCTGCTGAAGTGCGGCAAGGAGGGCCACCAGATGAGGACTGCACCGAGCGC 210
            || || ||| || || | || || || || || | || || || |
Db   160 GACCAGCGGCGGGGCCTCGAGCTTCCAGTTCCCGCCGCGGACCACCGGCACCGGC 219

Qy   211 CAGGCCAACTTCTTCCGCAGGACCTGGCCTTCCCCAGGACCAAGGCCCGGAGTTCCCC 270
            ||| || || | |||| |||| || ||||| |||| || | || |||
Db   220 TACATCCCGCCACCTCCATCCGACGCAGCGCCACCCGGATGCGGCTCAAGTCGCGGGGAG 279

Qy   271 AGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGAGCTGCGAGGTGCGCGGCGACAAC 330
            ||| | || | | | | || ||| | |||||| | || || || |
Db   280 CTCTACACCAACCGCGCCACGGGCGCCCTGCGACGACCACCACCGCCTCTGGACGTCGCAC 339

Qy   331 CCCGCAGCGGAGGCGGGCGCCAGGCGCCAGGGCACCCTGAACTTCCCCCAGATCACCCTG 390
            ||| | || | | || | |||| || ||||| |||| ||| | | |||
Db   340 GCCTTCGCGCACCGTCCACGCGTCTCGCGCTGCGCGGCGGCTGCGCCGGCCCGGCGGACGG 399

Qy   391 TGGCAGGCGCCCCTGGTGAGCATCAAGTGGGCGGCCAGATCAAGAGGAGGCCTGCTGGAC 450
            ||| || ||| || || | | || |||||| || || ||||| |||||||
Db   400 GTCGCGCGCCGACCTGCTGCGCTTCGCGTCGCGCGGCCGACCACTACGCCGTCGGCGCGGC 459

Qy   451 ACCGGCGCCACGACACCGCTGCTGGAGGAGATGAGCCTGCCGCGCGCAGTGGAAGCCCAAG 510
            || ||| ||| | | | | || | ||| || || | ||| ||||||||||
Db   460 GCCTGGAAGCGGGGGTACCTGCTCCACG------GCCGCCCGGCACGGCGCACGACCAGC 513

Qy   511 ATGATCGGCGCCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAG 570
            || | || |||| | || || || | | || | || | || | ||
Db   514 CTCGTCGCCGCTGCCATCGGCGTCGGCTTCGACGTCTCACGACCTCGGAGCTGACCACCACG 573

Qy   571 ATCTGCGGCAAGAGAGCCATCGGCACCGTGCTGATCGGCCGTCTCCACCCCGTGAACATCATC 630
            || | ||| || || || || | | | ||| ||| ||| | | ||| | |
Db   574 GTGCCCACCAACTCCCACCTCCGCGCGCCTGTGATCGTCGTCTCCACCACGCGCCCAAGTCCGTCGTC 633

Qy   631 GGCGGCAA 638
            ||| ||
Db   634 GTCGTCGA 641

RESULT 14
BE601575
LOCUS   HVSMEh0098K08f Hordeum vulgare 5-45 DAP spike EST library
             HVCDNA0009 (5 to 45 DAP) Hordeum vulgare subsp. vulgare cDNA clone
             HVSMEh0098K08f, mRNA sequence.
             640 bp   mRNA   linear   EST 22-OCT-2001
ACCESSION   BE601575
VERSION   BE601575.3 GI:16322423
KEYWORDS   EST.
SOURCE   Hordeum vulgare subsp. vulgare
  ORGANISM   Hordeum vulgare subsp. vulgare
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Pooideae; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 640)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D.,
             Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,
             Fenton,R.D., Close,S.J., Oates,R. and Main,D.
  TITLE   Development of a genetically and physically anchored EST resource
             for barley genomics: Morex 5-45 DAP spike cDNA library
  JOURNAL   Unpublished (2001)
COMMENT   On Aug 21, 2000 this sequence version replaced gi:13190104.
             Contact: Wing RA
             Clemson University Genomics Institute
             Clemson University
             100 Jordan Hall, Clemson, SC 29634, USA
             Tel: 864 656 7288
             Fax: 864 656 4293
             Email: rwing@clemson.edu
             Total hq bases = 461
             Seq primer: AATTAACCCTCACTAAAGGG
             High quality sequence stop: 534.
FEATURES   Location/Qualifiers
     source   1..640
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="Morex"
                     /sub_species="vulgare"
                     /db_xref="taxon:112509"
                     /clone="HVSMEh0098K08f"
                     /tissue_type="5-45 DAP Spike"
                     /clone_lib="Hordeum vulgare 5-45 DAP spike EST library
                     HVCDNA0009 (5 to 45 DAP)"
                     /note="Vector: lambdaZAP; Site1; EcoR1; Site 2: XhoI;
                     Plants were grown in the greenhouse at the University of
                     California, Riverside (Fenton, SJ Close, TJ Close). Whole
                     spikes with awns trimmed were collected at 5, 10, 15, 20,
                     30 and 45 DAP (Fenton). Total RNA was prepared from each
                     pool, equal quantities of all six RNA pools were combined,

FEATURES
     source   1..688
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="OSJNEb12C01"
                     /tissue_type="Leaf"
                     /dev_stage="3 week"
                     /lab_host="DH10B"
                     /clone_lib="OSJNEb"
                     /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                     XhoI; 24 hrs after innoculation with Rice Blast (Che
                     86061)"

ORIGIN

Query Match   2.9%; Score 72.2; DB 14; Length 688;
Best Local Similarity 46.2%; Pred. No. 26;
Matches 276; Conservative   0; Mismatches 318; Indels   0; Gaps   3;

Qy   107 GCGGCAAGGAGGGCCACATGCCCGCGCCCCCGCAAGAGGGCTGCTGGA   166
          |||| |||||| || |||||| |||| || |||||||||| || ||
Db   92 GCGCGGCGGCGCCCGAGATGCCGGCGCCACCTCGCAGAGGTGAAGCGGCGCATGGTGGCGC 151

Qy   167 AGTGCGGCAAGGAGGAGGGCCACCAGATGAAGGACTGCAGCGGCGCCGGCGGCCAACTTCTTCC 226
          || |||| || |||| ||||| |||||||||| |||
Db   152 AGTTCCGGCGAGCCGTTGGTCAAGATCCAGGGCGCACCCTTCGCGGTCGGGCCAGGTGGCCG 211

Qy   227 GCGAGGACCTGGCCTTCCCCCAGGGCAAGGCCCGGCCCGGCGAGTTCCCCAGCGAGCAGAACCGCG 286
          ||||||| |||| || |||| || | | || | ||| || | ||||||||
Db   212 CCGTCGCCCAGGCCAGGCCAGGCGGCGCCGGCGTGCCGCGTGTCCGTCTAGCTGCAGGAGGAGGGCCGCC 271

Qy   287 CCAACAGCCCACCAGCCCGCGAGCTGCAGGTGCGGCGGCGACAACCCCGCAAGCGAGGGCCG 346
          |||| || |||| ||| |||| || ||||| || |||| || ||| ||||
Db   272 CCCGCGTCAAGGCCAGCAGCGAGTGGATCCTCAACTGCATCGCCCACCGGCGCGGCGACATCT 331

Qy   347 GCGGCCAGCGGCGCAGGCGGCCGGCGCACCCTGGCGCCCCCCCTGG   406
          ||||| || | | |||||| | || ||||| || || | ||
Db   332 ACGGCGTCACCACCGGCTTCGGCGGCACCTCCACCCGCACCAAGGACGGCCCCGCCCGCGCCC 391

Qy   407 TGAGCATCAAGGTGGGCGGCCAGATCAAGCGGCATCTCAACGCCGGAATCTTCGGCACTGGCCCGATGGCC 463
          | |||||||| ||| ||||||||| ||||| | | |||||
Db   392 TCCAAGTCGAGCTCCTCAGGCATCTCAACGGCCGGAATCTTCGGGCACTCGGCACTGGCTCGATGGCC 451

Qy   464 ACACCGTGTCGAGGAGGAGGAGCCTGCCCGCCAAGTCGAAGCCCAAGATGATCGGGGCCA 523
          |||||||| |||||| || | ||||| || ||| | | || |||| ||||| ||
Db   452 ACACGGCTTCCGTCGGCCGACGGTTGCGGCGGCATCGTCGTGCGCATCAAACCCCTCCTCC 511

Qy   524 TCGGCGGCTTCATCAAGGTGCCCAGTACGACCAGAGATCCTGATCGAGATCCTCGGGCAAGA 583
          || ||| || || |||| | ||| | ||| || || || | |||| ||| |||||
Db   512 AGGGCTACTCCCGCATCCGGTTCGAGATCCTCGAGGCCATCACCACCAAGGCTGCTCAACACCG 571

Qy   584 AGGCCATCGGCACCGTGCTGATCGGCCCACCCCCGTGAACATCATCGGCCCGCCAACATGC 643
          ||| ||
Db   572 GCGTCACCGCCGTCCTCGGCCTGTGCCACCATCACCGCGGTCCGGTGACCTGGTTCCCC 631

ORIGIN

Query Match   2.9%; Score 72.2; DB 10; Length 640;
Best Local Similarity 47.7%; Pred. No. 25;
Matches 209; Conservative   0; Mismatches 229; Indels   0; Gaps   0;

Qy   789 GGAGATGAGAAGGAGGGCCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCC 848
          |||| | || |||| |||| | |||| ||||| | | || |
Db   110 GGAGGAGGTGACGGTGAAGGTGTCGGCCAAGATGATCGGTGACGGGGCCGGCGGCGCAC 169

Qy   849 CGTGTTGCCCATCAAGAAGAAGAGGACAGCACCACCAAGTAGCCAAGATGGCGCAAGTTCCGGCA 908
          | | ||| ||||| |||| | | | | | || || || | ||||
Db   170 CCTGACCCGCAACTTCAAGCACCTCAACCTCGACTTCCAGCTGCAGGACGGCGGGGCCGCAA 229

Qy   909 GCTGAACAAGCGCCACCCAGGACTTCTGGGAGGTGCAGCTGGCAGTCCCCACCCCGCCGG 968
          ||| || || || |||| ||| | | || |||| ||||| | || |||| || |
Db   230 GCTCAAGGTGGACGCCTGGTTCGGCACCGTCGGCGCCGCCATCGGCGACACCGCCCCAT 289

Qy   969 CCTGAAGAGAGAAGAGAGCCTGACGTGCTGGAAGCGTGGGGCGACGCCTACTTCAGCGTGCC 1028
          | | |||| || | | ||| | || || ||| |||| ||| | | | |||| |
Db   290 CTCCCACGTCCAGAACCTCATCACGGCGGCGTCACCAAGGGCTTCCGCTACACAAGATGGCCTT 349

Qy   1029 CCTGGACGAGGACTTCCGGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGAC 1088
          | ||| | || || | | ||| ||| || || | || | ||| ||||| | ||
Db   350 CGTCTACGCTCACTTCCCCATCAACGCCTTCCATCACCGGGCCATCGGGGGCATCGAGAT 409

Qy   1089 CCCCGGCATCGCGCTACCAGTACAACGTGTCGCCCAGGGTGAAGGGCAGCCCCAGCAT 1148
          || || ||| || || || |||| ||| | || || ||| | || ||
Db   410 CCGCAACTTCCTCCGGCGAGGAAGAAGGTGAGGAAGGTGAGGACATGCTCGACGGGGTCACCAT 469

Qy   1149 CTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCCGAACCCCGAGATCGT 1208
          ||| || | || || || |||||| | || || | | | || |||||| || ||
Db   470 CTTGCGGTCCGAGAAGGTCAAGGATGAGATCGTTCCTCACGCGCAAGCGAAGGGCTCTGT 529

Qy   1209 GATCTACCAGGCCCCCCT 1226
          | || | | | |||
Db   530 NTTCCGGTCCGGCGCCCT 547

RESULT 15
CB648640
CB648640

LOCUS   CB648640   688 bp   mRNA   linear   EST 08-APR-2003
DEFINITION OSJNEb12C01.f OSJNEb Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEb12C01 5', mRNA sequence.
ACCESSION   CB648640
VERSION   CB648640.1 GI:29643633
KEYWORDS   EST.
SOURCE   Oryza sativa (japonica cultivar-group)
  ORGANISM Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 688)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,

us-09-610-313-31.rst

```
Qy      644  TGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCG  700
             ||||||||||| ||||| |||||| |||   ||| | |||||| |||||  ||||||
Db      632  TGTCCTACATTGCCGGCCTCATCACCGGCCGCCCCAACGCGCAGGCCATCTCGCCCG  688
```

Search completed: April 10, 2004, 16:16:24
Job time : 4143 secs

OM nucleic - nucleic search, using sw model

Copyright   GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

Run on: April 10, 2004, 02:53:16 ; Search time 6326.67 Seconds
                                            (without alignments)
                                            16873.640 Million cell updates/sec

Title: US-09-610-313-31
Perfect score: 2463
Sequence: 2463
1 gtcgacgccaccatggccga.........gggctagcaccggtgaattc 2463

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151699 5 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

SUMMARIES

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2463	100.0	2463	6	AX455915	AX455915 Sequence
2	2442.2	99.2	2469	6	AX455914	AX455914 Sequence
3	2436.2	98.9	2457	6	AX455916	AX455916 Sequence
4	2046	83.1	2306	6	BD263704	BD263704 Improved
5	2046	83.1	2306	6	AR373387	AR373387 Sequence
6	2027.6	82.3	9166	6	AX427930	AX427930 Sequence
7	2025.2	82.2	2312	6	BD263706	BD263706 Improved
8	2025.2	82.2	2312	6	AR373389	AR373389 Sequence
9	2019.2	82.0	2300	6	BD263705	BD263705 Improved
10	2019.2	82.0	2300	6	AR373388	AR373388 Sequence
11	2019.2	82.0	9788	6	AX427936	AX427936 Sequence
12	2005.2	81.4	9169	6	AX427931	AX427931 Sequence
13	2001.8	81.3	9194	6	AX427926	AX427926 Sequence
14	2000.2	81.2	9194	6	AX427925	AX427925 Sequence
15	2000.2	81.2	12411	6	AX427927	AX427927 Sequence
16	1996.8	81.1	9785	6	AX427938	AX427938 Sequence
17	1993.4	80.9	9167	6	AX427933	AX427933 Sequence
18	1993.4	80.9	9170	6	AX427928	AX427928 Sequence
19	1993.4	80.9	9782	6	AX427935	AX427935 Sequence
20	1993.4	80.9	9783	6	AX427934	AX427934 Sequence
21	1993.4	80.9	9792	6	AX427932	AX427932 Sequence
22	1991.8	80.9	9189	6	AX427921	AX427921 Sequence
23	1988.0	80.7	4353	12	AF287353	AF287353 Synthetic
24	1972.2	80.1	3015	6	AX455946	AX455946 Sequence
25	1955.6	79.4	3009	6	AX455987	AX455987 Sequence
26	1942.6	78.9	4319	6	BD263639	BD263639 Improved
27	1942.6	78.9	4319	6	AR373322	AR373322 Sequence
28	1932	78.4	4352	12	AF287352	AF287352 Synthetic
29	1886.6	76.6	2577	6	AX457088	AX457088 Sequence
30	1879	76.3	4343	6	AX306428	AX306428 Sequence
31	1878.8	76.3	2305	6	BD263702	BD263702 Improved
32	1878.8	76.3	2305	6	AR373385	AR373385 Sequence
33	1875	76.1	4341	6	AX306429	AX306429 Sequence
34	1865.4	75.7	1977	6	AX455954	AX455954 Sequence
35	1857.8	75.4	1965	6	AX455952	AX455952 Sequence
36	1852.4	75.2	7897	6	AX427923	AX427923 Sequence
37	1852	75.2	2299	6	BD263703	BD263703 Improved
38	1852	75.2	2299	6	AR373386	AR373386 Sequence
39	1798.8	73.0	3024	12	AF287355	AF287355 Synthetic
40	1784	72.4	9407	6	AX427937	AX427937 Sequence
41	1651.8	67.1	4307	6	AR302563	AR302563 Sequence
42	1651.8	67.1	4307	6	AX019132	AX019132 Sequence
43	1651.8	67.1	4307	6	BD130468	BD130468 Antiviral
44	1632.6	66.3	4307	6	BD268901	BD268901 Anti-vira
45	1632.6	66.3	4307	6	AX035453	AX035453 Sequence

ALIGNMENTS

RESULT 1
AX455915
LOCUS   AX455915   2463 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION   Sequence 31 from Patent WO0204493.
ACCESSION   AX455915
VERSION   AX455915.1 GI:21714900
KEYWORDS
SOURCE   synthetic construct
ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE   Polynucleotides encoding antigenic hiv type c polypeptides,
             polypeptides and uses thereof
JOURNAL   Patent: WO 0204493-A 31 17-JAN-2002;

Pred. No. is the number of results predicted by chance to have a

```
            CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES    Location/Qualifiers
source      1..2463
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="PR975YM"
ORIGIN

Query Match          100.0%;   Score 2463;   DB 6;   Length 2463;
Best Local Similarity 100.0%;   Pred. No. 5.7e-249;
Matches 2463;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;

Qy      1 GTCGACGCCACCATGGCCGAGGCCATGAGCGAGGCCAGCCAGCGGCCACATCCTGATGCAG    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCGACGCCACCATGGCCGAGGCCATGAGCGAGGCCAGCCAGCGGCCACATCCTGATGCAG    60

Qy     61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGC   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGC   120

Qy    121 CACATCGCCCGCAACTGCGGCGGCCCCGCCCCCGCCCCCGGAAGAAGGGTGCTGGAAGAG   180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACATCGCCCGCAACTGCGGCGGCCCCGCCCCCGCCCCCGGAAGAAGGGTGCTGGAAGAG   180

Qy    181 GGCCACCAGATGAAGGACTGCACCGAGCGGCCAGGCGCCAACTTCTTCCGCGAGGACCTGCC   240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCCACCAGATGAAGGACTGCACCGAGCGGCCAGGCGCCAACTTCTTCCGCGAGGACCTGCC   240

Qy    241 TTCCCCAGGGCAAGGCCAAGGCCCGCGAGTTCCCGCAGCGCAGAACCGGCGCCAACAGCCCACC   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCCCCAGGGCAAGGCCAAGGCCCGCGAGTTCCCGCAGCGCAGAACCGGCGCCAACAGCCCACC   300

Qy    301 AGCCGCGAGTCGCAGGTGCGCGGCGGACAACCCCCGCAGGGAGGCCGGCGCCGAGCGCCAG   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCCGCGAGTCGCAGGTGCGCGGCGGACAACCCCCGCAGGGAGGCCGGCGCCGAGCGCCAG   360

Qy    361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCTGTGACGTCATCAAGGTG   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCTGTGACGTCATCAAGGTG   420

Qy    421 GGCGGCCAGTACGACCAGATCCTGATCGAGAATCTGTGGGCAGGAAGGCCATCGGCACCGTG   480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGCGGCCAGTACGACCAGATCCTGATCGAGAATCTGTGGGCAGGAAGGCCATCGGCACCGTG   480

Qy    481 ATGAGCTGCCGCGGCAAGTGAAGCCCAAGATGATCGGCCAAGATGATCGGCCCAAGAGTGGAAGTC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGAGCTGCCGCGGCAAGTGAAGCCCAAGATGATCGGCCAAGATGATCGGCCCAAGAGTGGAAGTC   540

Qy    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGGGCGCCAAGAAGGCCATCGGCACCGTG   600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGGGCGCCAAGAAGGCCATCGGCACCGTG   600

Qy    601 CTGATCGGCCCACCCCACCCCCCGTGAACATCATCGGCCGCAACATCATCGGCCTGGGCTGC   660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGATCGGCCCACCCCACCCCCCGTGAACATCATCGGCCGCAACATCATCGGCCTGGGCTGC   660

Qy    661 ACCTGACTTCCCCATCAGCGCCCATCAGCGTGCCCGTGAAGTCTGAAGCCCGGCCGGCCATG   720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCTGACTTCCCCATCAGCGCCCATCAGCGTGCCCGTGAAGTCTGAAGCCCGGCCGGCCATG   720

Qy    721 GACGGGCCCCAAGGTGAAGCAGTGCTGCCCCTGACGGAGAAGATCAAGGCCCCTGACCGCC   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACGGGCCCCAAGGTGAAGCAGTGCTGCCCCTGACGGAGAAGATCAAGGCCCCTGACCGCC   780

Qy    781 ATCTGGGAGGAGATGGAGAAGGGCAAGATCATCGGCCAGATCACCACCAAGAACCCCTAC   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ATCTGGGAGGAGATGGAGAAGGGCAAGATCATCGGCCAGATCACCACCAAGAACCCCTAC   840

Qy    841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC   900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC   900

Qy    901 TTCCCGGAGCTGAACAAGCGCACCAGGCGACTTCTCGGAGCTGCAGCGTGCAGCTGGGCATCCCCAC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTCCCGGAGCTGAACAAGCGCACCAGGCGACTTCTCGGAGCTGCAGCGTGCAGCTGGGCATCCCCAC   960

Qy    961 CCCGGCCGGCCTGAAGAAGAGAAGGAGCGTGACCGTGCTGACCGGTGGGCGACGCGCTTACTTC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCCGGCCGGCCTGAAGAAGAGAAGGAGCGTGACCGTGCTGACCGGTGGGCGACGCGCTTACTTC  1020

Qy   1021 AGCGTGCCCCTGGACGAGGACTTCGCCAAGCTACACCGCCTTCACCATCCCAGCATCAAC  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGCGTGCCCCTGGACGAGGACTTCGCCAAGCTACACCGCCTTCACCATCCCAGCATCAAC  1080

Qy   1081 AACGAGACCCCGGCATCCGGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACGAGACCCCGGCATCCGGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140

Qy   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCC  1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCC  1200

Qy   1201 GAGATCGTGATCTACCAGGGCCCCCGTACGTGGGCAGCGACTCGGCAGCAC  1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCAGGGCCCCCGTACGTGGGCAGCGACTCGGCAGCAC  1260

Qy   1261 CGCGCCAAGATGAGGAGCTGCGCAAGCACCTGCGCTGCGAGGACCTGCC  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGCCAAGATGAGGAGCTGCGCAAGCACCTGCGCTGCGAGGACCTGCC  1320

Qy   1321 AAGAAGCACCAGAAGGAGCCCCCTTCGTGGGATGGGCTACGAGCTGCACCGACAAAG  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGAAGCACCAGAAGGAGCCCCCTTCGTGGGATGGGCTACGAGCTGCACCGACAAAG  1380

Qy   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGAACCTGGACATCCAG  1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGAACCTGGACATCCAG  1440

Qy   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGATCCACCCCCGGCAGTCAAGGTG  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGATCCACCCCCGGCAGTCAAGGTG  1500

Qy   1501 CTGTGCAAGCTGCTGCGGCGGCCCGAGAGCATCGTCCCCTGACCGAGGAG  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGGCGGCCCGAGAGCATCGTCCCCTGACCGAGGAG  1560

Qy   1561 GCCGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGGCGTGTAC  1620
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGGCGTGTAC  1620

Qy   1621 TACCACCCCAGCAAGGACCTGGTGCCGAGAACCAGAGTTCCGCCAGTTGGACC  1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TACCACCCCAGCAAGGACCTGGTGCCGAGAACCAGAGTTCCGCCAGTTGGACC  1680

Qy   1681 TACCAGATCTACCAGGAGACCTTCAAGAACCTGAAGACCGGCAAGATGGC  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACCAGATCTACCAGGAGACCTTCAAGAACCTGAAGACCGGCAAGATGGC  1740

Qy   1741 ACGCCCACCACCAACGACGTGAACGACGTGACCGGCCGTGACCCCATGGAG  1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACGCCCACCACCAACGACGTGAACGACGTGACCGGCCGTGACCCCATGGAG  1800

Qy   1801 AGCATCTGATCTGGGCAAGACCCCAAGTTCCGCCTGCCCAGTTCCGCCTGG  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCTGATCTGGGCAAGACCCCAAGTTCCGCCTGCCCAGTTCCGCCTGG  1860

Qy   1861 GAGACTGGTGACCGACTACTGGCAGGCCCACTGGAGTTCGTGAAC  1920
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGACTGGTGACCGACTACTGGCAGGCCCACTGGAGTTCGTGAAC  1920

Qy   1921 ACCCCCCCGTGGTGAAGCTGTGGTACCAGCTGGGCCAAGCTGGTGGAC  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACCCCCCCGTGGTGAAGCTGTGGTACCAGCTGGGCCAAGCTGGTGGAC  1980

Qy   1981 ACCTTCTACGTGACGGCGCCCCAACCGCGACCAAGATCGGCAGGCCGGCTACGTG  2040
```

```
Db   1981 ||||||||||||||||||||||||||||||||||||||||
          ACCTTCTACGTGGACGGCGCCGCCAACCGGCGAGACCAAGATCGGCAAGGCGGCTACGTG 2040
Qy   2041 ACCGACCGGCGGGCCGGGCCGGCGAGAAGATCGTGAGCCTGAGCGAGACCAGAGACCGAG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACCGGCGGGCCGGGCCGGCGAGAAGATCGTGAGCCTGAGCGAGACCAGAGACCGAG 2100
Qy   2101 CTGCAGGCCATCCAGTGGCCCTGAGGACACGCGGCAGCGGCAGCGGGTGAACATCGTGACCGAC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGCCATCCAGTGGCCCTGAGGACACGCGGCAGCGGCAGCGGGTGAACATCGTGACCGAC 2160
Qy   2161 AGCCAGTACCGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGCGGAGAGCGAGCTGGTG 2220
          |||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACCGCCCTGGGCATCATCCAGGCCCAGCCCGACAACCCCGAGCGGAGCTGGTG 2220
Qy   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAAGGTGTACCTGAGCTGGTGCCCGCC 2280
          |||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAAGGTGTACCTGAGCTGGTGCCCGCC 2280
Qy   2281 CACAAGGGCATCGGCGCGCGGCGGCGGCCAACGAGCAGATCGACAAGCTGGTGACGACCTG 2340
          |||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCGCGCGGCGGCGGCCAACGAGCAGATCGACAAGCTGGTGACGACCTG 2340
Qy   2341 GTGCTGTTCCTTGGACGGCGGCGGCCCCTTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2400
          |||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTTGGACGGCGGCGGCCCCTTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2400
Qy   2401 TACGTGGCAGGCGGCGGCCGGCGCCCCTTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460
          |||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGCAGGCGGCGGCCGGCGCCCCTTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460
Qy   2461 TTC 2463
          |||
Db   2461 TTC 2463
```

```
Query Match       99.2%;  Score 2442.2;  DB 6;  Length 2469;
Best Local Similarity   99.6%;  Pred. No. 8.6e-247;
Matches 2460;  Conservative   0;  Mismatches   3;  Indels   6;  Gaps   1;

Qy    1 GTGCACGCGCCACCATGGCCGGCCATGAGCCCAGCGGCCAACATCCTGATGCAG 60
         ||||||||||||||||||||||||||||||||||||||||||
Db    1 GTGCACGCGCCACCATGGCCGGCCATGAGCCCAGCGGCCAACATCCTGATGCAG 60
Qy   61 CGGCAGCAACTTCAAGGGCCCAAGGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120
         ||||||||||||||||||||||||||||||||||||||||
Db   61 CGCAGCAACTTCAAGGGCCCAAGGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120
Qy  121 CACATCGCCCGCAACTGCCGCGCCCCCCGCGGGAGGGCTGCTCGGAAGTGCGGCAAGGAG 180
```

```
Db  121 CACATCGCCCGCAACTGCCGCGCCCCCCGCGGGAGGGCTGCTGGAAGTGCGGCAAGGAG 180
Qy  181 GGCCACCAGCAGATGAAGGACTGCGACCTGAGCGGGCCGAGCTGAGACTTCTTCCGCGAGGACCTGCC 240
         |||||||||||||||||||||||||||||||||||||
Db  181 GGCCACCAGCAGATGAAGGACTGCGACCTGAGCGGGCCGAGCTGAGACTTCTTCCGCGAGGACCTGCC 240
Qy  241 TTCCCCAGGGCAAGGCCCGAGTTCCCCAGCGAGACAGATCGTGACCGAC 300
         |||||||||||||||||||||||||||||||||
Db  241 TTCCCCAGGGCAAGGCCCGAGTTCCCCAGCGAGCAGCAACAGCCCCACC 300
Qy  301 AGCCGGGAGCTGCAGTGGCGGGCGGCGACAACCCCCGCAGCCGCGAGGCGCGGCAGCCCAG 360
         |||||||||||||||||||||||||||||||||
Db  301 AGCCGGGAGCTGCAGTGGCGGGCGGCGACAACCCCCGCAGCCGCGAGGCGCGGCAGCCCAG 360
Qy  361 GGCACCCTGAACTTCCCCAGGCTGTGGCAGGCCCCTGGTGCAGGCATCAAGTG 420
         ||||||||||||||||||||||||||||||||
Db  361 GGCACCCTGAACTTCCCCAGGCTGTGGCAGGCCCCTGGTGCAGGCATCAAGTG 420
Qy  421 GGCGCGCCAGATCAAGGAGCCCTGCTGACACCGGCGCGCGGCATCGGCGGCGCTTCATCAAG 480
         |||||||||||||||||||||||||||||||
Db  421 GGCGCGCCAGATCAAGGAGCCCTGCTGACACCGGCGCGCGGCATCGGCGGCGCTTCATCAAG 480
Qy  481 ATGAGCCTGCCCGGCAAGTGGAGCCCAAGATGATCGGCGGCATCTGCGCGAGATCCTGATCGA 540
         ||||||||||||||||||||||||||||||
Db  481 ATGAGCCTGCCCGGCAAGTGGAGCCCAAGATGATCGGCGGCATCTGCGCGAGATCCTGATCGA 540
Qy  541 GTGGGCCAGTACGACCAGATCCTGATCGGGCGGCTAGCACCGGTGAA 600
         ||||||||||||||||||||||||||||
Db  541 GTGGGCCAGTACGACCAGATCCTGATCGGGCGGCAAGAGGCCATCGGGCCAACATGCTGGGCGTG 600
Qy  601 CTGATCGGCCCACCCCGTGACCATCAGCCCCATCGACCCGTGAAGCTGAAGCCCGTGACCGCCGCC 660
         ||||||||||||||||||||||||||||
Db  601 CTGATCGGCCCACCCCGTGACCATCAGCCCCATCGACCCGTGAAGCTGAAGCCCGTGACCGCC 660
Qy  661 ACCCTGAACTTCCCCATCAGCCCATCAGCCCATCGACCCGTGAAGCTGAAGGCCCTGACCGCC 720
         ||||||||||||||||||||||||||||
Db  661 ACCCTGAACTTCCCCATCAGCCCATCAGCCCATCGACCCGTGAAGCTGAAGGCCCTGACCGCC 720
Qy  721 GACGGGCCCAAGGTGAAGCAGTGAGCAGTTCCTGACCGGAGAGATCAAGGCCCTGACCGAC 780
         ||||||||||||||||||||||||||||||
Db  721 GACGGGCCCAAGGTGAAGCAGTGAGCAGTGGCCCTGACCGGAGAGATCAAGGCCCTGACCGAC 780
Qy  781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840
         ||||||||||||||||||||||||||||||
Db  781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840
Qy  841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGTGGAC 900
         ||||||||||||||||||||||||||||||
Db  841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
Qy  901 TTCCGGCGAGCTGAACAAGCGACACCCAGGACTTCTGGGAGGTGCAGCTGCAGCTGGGCATCCCCCAC 960
         ||||||||||||||||||||||||||||||
Db  901 TTCCGGCGAGCTGAACAAGCGACACCCAGGACTTCTGGGAGGTGCAGCTGCAGCTGGGCATCCCCCAC 960
Qy  961 CCGGCGCCTGAAGAAGAAGAAGCGTGACCGTGCTGACGTGGGCGACGTGGGCGACGCCTACTTC 1020
         ||||||||||||||||||||||||||||||
Db  961 CCGGCGCCTGAAGAAGAAGAAGCGTGACCGTGCTGACGTGGGCGACGTGGGCGACGCCTACTTC 1020
Qy 1021 AGCGTGCCCTGGACGAGGACTTCCGGACCTACACCCGCCTTCACCATCCCCAGCGATCAAC 1080
         ||||||||||||||||||||||||||||||
Db 1021 AGCGTGCCCTGGACGAGGACTTCCGGACCTACACCCGCCTTCACCATCCCCAGCGATCAAC 1080
Qy 1081 AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
         ||||||||||||||||||||||||||||||
Db 1081 AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
Qy 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCC 1200
         ||||||||||||||||||||||||||||||
Db 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCC 1200
Qy 1201 GAGATCGTGATCTACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGC 1254
         ||||||||||||||            ||   ||||||||||||||||
```

```
Db  1201 GAGATCGTGATCGTTACCAGTACATGGACGACCTGACGTGGGCAGCGACCTGGAGATCGGC 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1255 CAGCACCGGCGCCAAGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACC 1314
Db  1261 CAGCACCGGCGCCAAGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACC 1320

Qy  1315 CCCGACAAGAGCACCAGAGAGGAGCCCCCCTTCCTGTGATGGGCTACGAGCTGCACCCC 1374
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CCCGACAACAAGAAGCACCAGAGAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCC 1380

Qy  1375 GACAAGTGACCGTGCAGCCCATCGAGCTGCCCCAGAAGGAGAGCTGACCGTGAACGAC 1434
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCCAGAAGGAGAGCTGACCGTGAACGAC 1440

Qy  1435 ATCCAGAAGCTGGTGGTGGGCAAGCTGAACTGGGCCAGATCTACCCCGGCATCAAGTG 1494
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 ATCCAGAAGCTGGTGGTGGGCAAGCTGAACTGGGCCAGATCTACCCCGGCATCAAGTG 1500

Qy  1495 CGCCAGCTGTGTGCAAGCTGCTGCGCGGCGCGCAAGGCCCTGACCGACATCGTGCCCCTGACC 1554
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CGCCAGCTGTGTGCAAGCTGCTGCGCGGCGCGCAAGGCCCTGACCGACATCGTGCCCCTGACC 1560

Qy  1555 GAGGAGGCCGAGCTGGAGCTGGCCGAGAGATCCTGCCGAGAACCGCTGCCGAGCCCGTGCACGGC 1614
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GAGGAGGCCGAGCTGGAGCTGGCCGAGAGATCCTGCCGAGAACCGCTGCCGAGCCCGTGCACGGC 1620

Qy  1615 GTGTACTACGACCCCGACCAAGGACCTGGTGCGCCAGAGATCCAGAGAGCAGGGCCACGACCAG 1674
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 GTGTACTACGACCCCGACCAAGGACCTGGTGCGCCAGAGATCCAGAGAGCAGGGCCACGACCAG 1680

Qy  1675 TGGACCTACCAGATCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1734
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGGACCTACCAGATCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1740

Qy  1735 ATGGCCACCGCCCACAACAACCACCACGACGTGAAGCAGCTGAGCAGCGGCCGTGCGAGAGATCGCC 1794
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ATGGCCACCGCCCACAACAACCACCACGACGTGAAGCAGCTGAGCAGCGGCCGTGCGAGAGATCGCC 1800

Qy  1795 ATGGAGAGCATCGTGATCGTTCGGGGCAAGACCCCCAAGTTCCGCTGCCATCCAGAAGGAG 1854
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 ATGGAGAGCATCGTGATCGTTCGGGGCAAGACCCCCAAGTTCCGCTGCCATCCAGAAGGAG 1860

Qy  1855 ACCTGGGAGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC 1914
              |||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACCTGGGAGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC 1920

Qy  1915 CGCAGCAACTTCAAGGGCAACTGCCGCCCCCCAGCCATCATCCAAGTGCTTCAACTCGGC 1974
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CGCAGCAACTTCAAGGGCAACTGCCGCCCCCAGCCATCATCCAAGTGCTTCAACTCGGC 1980

Qy  1975 GCCGAGACCTTCTACGTGACGAGCCGGCCGGCCCGAGATCGGCAAGGCCGGC 2034
              |||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 GCCGAGACCTTCTACGTGGACGAGCCGGCCGGCCCGAGATCGGCAAGGCCGGC 2040

Qy  2035 TACGTGACCGAGCGGGGGCCGGCCGGGGGCCGCCAACCAGAAG 2094
              |||||||||||||||||||||||||||||||||||||||||
Db  2041 TACGTGACCGAGCGGGGGCCGGCCGGGGGCCGCCAACCAGAAG 2100

Qy  2095 ACCGAGCTGCAGGCCATCGAGCGGCCTGGCCCTGCAGGAGTGAACATCGTG 2154
              |||||||||||||||||||||||||||||||||||||||||||||
Db  2101 ACCGAGCTGCGAGGCCATCGAGCTGCGGCCTGGCCCTGCAGGAGTGAACATCGTG 2160

Qy  2155 ACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGGCGAGAGCGCAG 2214
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGGCGAGAGCGCAG 2220

Qy  2215 CTGGTGAACCAGATCATCGAGCAGCTGTACCTGAGCTGGGTG 2274
              |||||||||||||||||||||||||||||||||||||||||
Db  2221 CTGGTGAACCAGATCATCGAGCAGCTGTACCTGAGCTGGGTG 2280

Qy  2275 CCCGCCCAAGGGCCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATC 2334
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CCCGCCCCAACAAGGGCCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATC 2340
```

```
Qy  2335 CGCAAGGTGCTGTTCCTGACGACGGCATCGATGGCGGCATCGTGATCGATCTACCAGTACACATGGAC 2394
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 CGCAAGGTGCTGTTCCTGACGACGGCATCGATGGCGGCATCGTGATCGATCTACCAGTACACATGGAC 2400

Qy  2395 GACCTGCTACGTGGCCCAGCGGCCGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGCTAGCGACC 2454
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 GACCTGCTACGTGGCCCAGCGGCCGGCGGCGGCCCTAGGATCGATTAAAAAGCTTCCCGGGCTAGCACC 2460

Qy  2455 GGTGAATTC 2463
              ||||||||
Db  2461 GGTGAATTC 2469
```

RESULT 3
LOCUS   AX455916   2457 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION   Sequence 32 from Patent WO0204493.
ACCESSION   AX455916
VERSION   AX455916.1   GI:21714901
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
               artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE   Polynucleotides encoding antigenic hiv type c polypeptides,
               polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 32 17-JAN-2002;
               CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES   Location/Qualifiers
     source   1..2457
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PR975YMM"
ORIGIN

Query Match   98.9%;   Score 2436.2; DB 6; Length 2457;
Best Local Similarity 99.6%; Pred. No. 3.6e-246;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

```
Qy    1 GTCGACGCCACCATGGCCGAGGCCATGAGCAGCCAGCGCCAACATCCTGATGCAG 60
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GTCGACGCCACCATGGCCGAGGCCATGAGCAGCCAGCGCCAACATCCTGATGCAG 60

Qy   61 CGCAGCAACTTCAAGGGCAACTGCCGCCCCCCAGCCATCATCCAAGTGCTTCAACTCGGC 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGCAGCAACTTCAAGGGCAACTGCCGCCCCCAGCCATCATCCAAGTGCTTCAACTCGGC 120

Qy  121 CACATCGCCCGCAACTGCCCGCGCCCCCGCCCGCAAGAAGGCTGCTGGAAGTGCGGCAAGGAG 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 CACATCGCCCGCAACTGCCCGCGCCCCCGCCCGCAAGAAGGCTGCTGGAAGTGCGGCAAGGAG 180

Qy  181 GGCCACCAGATGAAGGACTGCACCGAGCGGCCCAGGCCCAACTTCTTCCGGCAAGGACCTGGCC 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 GGCCACCAGATGAAGGACTGCACCGAGCGGCCCAGGCCCAACTTCTTCCGGCAAGGACCTGGCC 240

Qy  241 TTCCCCAGGGACCAAGGCCCGCGAGTTCCCGGCACCCCCGCCGACACCCCCCACC 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TTCCCCAGGGACCAAGGCCCGCGAGTTCCCGGCACCCCCGCCGACACCCCCCACC 300

Qy  301 AGCCGCGAGCTGCAGGTGCGCGGCGGCGCAAACCCCAGCGGCCGGCCGGCCGCCAG 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AGCCGCGAGCTGCAGGTGCGCGGCGGCGCAAACCCCAGCGGCCGGCCGGCCGCCAG 360

Qy  361 GGCACCTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCCCTGGTGAGCATCAAGGTG 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GGCACCTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCCCTGGTGAGCATCAAGGTG 420

Qy  421 GGCGGCCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 GGCGGCCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG 480
```

```
Qy   481 ATGAGCCTGCCGGCAGTGAAGCCCAGAGATGATGGGCGGCATCGGCGGCTTCATCAAG 540
         ||||||||||||| |||| ||||||||||||||||  |||||||||||||||||||||||
Db   481 ATGAGCCTGCCGGCCAGTGGAAGCCCAGAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540

Qy   541 GTGCGGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGCGGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600

Qy   601 CTGATCGGCCCCACCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCCCCACCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC 660

Qy   661 ACCCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720

Qy   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC 780

Qy   781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTAC 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTAC 840

Qy   841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900

Qy   901 TTCCGCGAGCTGAACAAGAGGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCGCGAGCTGAACAAGAGGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC 960

Qy   961 CCCGCCGGCCTGAAGAAGAAGAAGAGCGTGCTGGACGTGCTGGACACCTACTTC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCCGGCCTGAAGAAGAAGAAGAGCGTGCTGGACGTGCTGGACACCCTACTTC 1020

Qy  1021 AGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAAC 1080

Qy  1081 AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140

Qy  1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCC 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCC 1200

Qy  1201 GAGATCGTGATCTACCAGGGCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GAGATCGTGATCTACCAGGGCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1260

Qy  1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1320

Qy  1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGCCCCTTCCTG----CAGCTGCACCCCGACAAG 1380
         |||||||||||||||||||||||||||||||||||           |||||||||||||||
Db  1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGT----CAGCTGCACCCCGACAAG 1374

Qy  1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1375 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAG 1434

Qy  1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAG 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1435 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAG 1494

Qy  1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1495 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1554


Qy  1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGATCCTGCGCGGGCGTGTAC 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1555 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGATCCTGCGCGGGCGTGTAC 1614

Qy  1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACC 1680
         ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  1615 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACC 1674

Qy  1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1675 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1734

Qy  1741 ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1735 ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1794

Qy  1801 AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1795 AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1854

Qy  1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1855 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1914

Qy  1921 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1915 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1974

Qy  1981 ACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1975 ACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2034

Qy  2041 ACCGACCGGGGCCGGCAGAAGATCGTGACCCTGACCGAGACCACCAACCAGAAGACCGAG 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2035 ACCGACCGGGGCCGGCAGAAGATCGTGACCCTGACCGAGACCACCAACCAGAAGACCGAG 2094

Qy  2101 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCAGGAGGTGAACATCGTGACCGAC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2095 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCAGGAGGTGAACATCGTGACCGAC 2154

Qy  2161 AGCCAGTACGCCCTGGGCATCATCCAGGCTGCGCCCGACAAGGAGGTGTACCTGGCCTGGGTG 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2155 AGCCAGTACGCCCTGGGCATCATCCAGGCTGCGCCCGACAAGGAGGTGTACCTGGCCTGGGTG 2214

Qy  2221 AACCAGATCATCGAGCAGCTGATCGAGAAGGAGAAGGTGTACCTGGCCTGGGTGCCCGCC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2215 AACCAGATCATCGAGCAGCTGATCGAGAAGGAGAAGGTGTACCTGGCCTGGGTGCCCGCC 2274

Qy  2281 CACAAGGGCATCGGCGGCAACGAGCAGGTCGACAAGCTGGTGAGCGCCGGCATCCGCAAG 2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2275 CACAAGGGCATCGGCGGCAACGAGCAGGTCGACAAGCTGGTGAGCGCCGGCATCCGCAAG 2334

Qy  2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGATGGCGGCATGGACGACCTG 2400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2335 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGATGGCGGCATGGACGACCTG 2394

Qy  2401 TACGTGGGCAGCGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2395 TACGTGGGCAGCGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2454

Qy  2461 TTC 2463
         |||
Db  2455 TTC 2457
```

```
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 2306)
  AUTHORS   Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
            Greer,C., Selby,M. and Walker,C.
  TITLE     Improved expression of HIV polypeptides and production of
            virus-like particles
  JOURNAL   Patent: JP 2002533124-A 71 08-OCT-2002;
            CHIRON CORP
COMMENT     OS   Artificial Sequence
            PN   JP 2002533124-A/71
            PD   08-OCT-2002
            PF   30-DEC-1999 JP 2000591193
            PR   31-DEC-1998 US    60/114495,01-DEC-1999 US    60/168471 PI
            SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
            KARIN HARTOG,
            PI   HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
            C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
            C12N5/10,
            PC   C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
            Description of Artificial Sequence: FS(-).protmod.RTopt.YM FH   Key
                      Location/Qualifiers
            FT   source          1. .2306
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .2306
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match            83.1%;  Score 2046;  DB 6;  Length 2306;
  Best Local Similarity  93.6%;  Pred. No. 2.4e-205;
  Matches 2159;  Conservative    0;  Mismatches  135;  Indels  12;  Gaps    2;

Qy      170 GCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCG 229
            |||||   || || ||||| ||||| || ||||| ||| | ||||| || ||||||||||
Db        1 GCGGCCGCGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTCTTCCGCG 60

Qy      230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCA 289
            |||||||||||||||  |||||||||||||||||||||  ||||||||||| ||||||||
Db       61 AGGACCTGGCCTTCCTGCAGGGCAAGGCCCGCGAGTTCAGCAGCGAGCAGACCCGCGCCA 120

Qy      290 ACAGCCCCACCAGCCGCGAGCTGCAGGTGCGCGGCG------ACAACCCCCGCAGCGAGG 343
            ||||||||||| ||||||||||||||||| | ||||      |||||  ||  |||||||
Db      121 ACAGCCCCACCCGCCGCGAGCTGCAGGTGTGGGGCGGCGAGAACAACAGCCTGAGCGAGG 180

Qy      344 CCGGCGCCGAGCGCCAGGGCACCCTG------AACTTCCCCCAGATCACCCTGTGGCAGC 397
            |||||||||| |||||||||||| ||      ||||||||||||||||||||||||||||
Db      181 CCGGCGCCGACCGCCAGGGCACCGTGAGCTTCAACTTCCCCCAGATCACCCTGTGGCAGC 240

Qy      398 GCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCG 457
            |||||||||||| ||||| | | ||||||||| |||||||||| ||||| ||||||||||
Db      241 GCCCCCTGGTGACCATCAGGATCGGCGGCCAGCTCAAGGAGGCGCTGCTCGACACCGGCG 300

Qy      458 CCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCG 517
            ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      301 CCGACGACACCGTGCTGGAGGAGATGAACCTGCCCGGCAAGTGGAAGCCCAAGATGATCG 360

Qy      518 GCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCG 577
            |||| ||||| ||||||||||||||||| ||||||||||||||||   | ||||||||||
Db      361 GCGGGATCGGGGGCTTCATCAAGGTGCGGCAGTACGACCAGATCCCCGTGGAGATCTGCG 420

Qy      578 GCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCA 637
            || | ||||||||||||||||||||| | |||||||||||||||||||||||||||||||
Db      421 GCCACAAGGCCATCGGCACCGTGCTGGTGGGCCCCACCCCCGTGAACATCATCGGCCGCA 480

Qy      638 ACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGC 697
            || ||||||||||| | |||||||||||||||||||||||||||||||||||||| ||||
Db      481 ACCTGCTGACCCAGATCGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACGGTGC 540

Qy      698 CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG 757
            |||||||||||||||| || ||||||||||||||||| ||||||||||||||||||||||
Db      541 CCGTGAAGCTGAAGCCGGGGATGGACGGCCCCAAGGTCAAGCAGTGGCCCCTGACCGAGG 600

Qy      758 AGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCA 817
            |||||||||||||||||   |  ||||||   ||||||||||||||||||||||||| ||
Db      601 AGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATCAGCA 660

Qy      818 AGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCA 877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 AGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCA 720

Qy      878 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 780

Qy      938 AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC 997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC 840

Qy      998 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCG 1057
            |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db      841 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACAAGGACTTCCGCAAGTACACCG 900

Qy     1058 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 1117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 960

Qy     1118 TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
            ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db      961 TGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020

Qy     1178 AGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCA 1237
            |||||||||||   |  |||||||| ||||||||||||||||||||||||||||||||||
Db     1021 AGCCCTTCCGCAAGCAGAACCCCGACATCGTGATCTACCAGGCCCCCCTGTACGTGGGCA 1080

Qy     1238 GCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1297
            |||||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
Db     1081 GCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGCCAGCACCTGCTGC 1140

Qy     1298 GCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGG 1357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGG 1200

Qy     1358 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGA 1417
            |||||||||||||||||||||||||||||||||||||||||  ||||||||||||||| |
Db     1201 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCATGCTGCCCGAGAAGGACA 1260

Qy     1418 GCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCT 1477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 GCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCT 1320

Qy     1478 ACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCG 1537
            || ||||||||||||||   |||||||||||||||||||||||| |||||||||||||||
Db     1321 ACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCG 1380

Qy     1538 ACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGC 1597
            |  |  | ||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db     1381 AGGTGATCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGA 1440

Qy     1598 GCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGA 1657
              |||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db     1441 AGGAGCCCGTGCACGAGGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGA 1500

Qy     1658 AGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGA 1717
            |||||||||| | |||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 AGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGA 1560

Qy     1718 CCGGCAAGTACGCCAAGATGCGCACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGG 1777
            ||||||||||||||   ||||||  |||||||||||||||||||||||||||||||||||
Db     1561 CCGGCAAGTACGCCCGCATGCGCGGCGCCCACACCAACGACGTGAAGCAGCTGACCGAGG 1620

Qy     1778 CCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCC 1837
            ||||||||||| |   ||  ||||||||||||||||||||||||| ||||||||||   |
```

Db 1621 CCGTGCGAGAAGGTGAGCACCGAGAGCATCGTGATCTGGGCAAGATCCCCAAGTTCAAGC 1680
             |||||
Qy   1 GCGGGCCGAGGACCAACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTCTTCCGCG 60

Qy 1838 TGCCCATCCAGAAGGAGACCTGGTGACCGACTACTGGCCAGGCCACCTGGA 1897
         ||||||||||||||||||||||||||||||||||||||||||||
Db   230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCGGCGAGTTCCCCGAGCGAGACCGGCGCCA 289

Db 1681 TGCCCATCCAGAAGGAGACCTGGTGGATGGAGTACTGGCCAGGCCACCTGGA 1740
         ||||||||||||||||||||||||||||||||||||||||||||
Qy   61 AGGACCTGGCCTTCCTGCAGGGCAAGGCCGGCGAGTTCAGCAGCGGAGACCGGCGCCA 120

Qy 1898 TCCCCGAGTGGGAGTTCGTGAACACACCCCCCGTGGTGAAGCTGTGGTACCAGTCTGGAGA 1957
         ||||||||||||||||||||||||||||||||||||||||||
Db   290 ACAGCCCCACCAGCGCCGAGCTGCAGGTGCGCGGGCG-----ACAACCCCGCAGCGGAGG 343

Db 1741 TCCCCGAGTGGGAGTTCGTGAACACACCCCCCGTGGTGAGCTGTGGTACCAGCTGGAGA 1800
         ||||||||||||||||||||||||||||||||||||||||
Qy   121 ACAGCCCCACCCACCCGCCGAGCTGCAGGTGTGGGCGGGCGGAGAACAACAGCCTGAGGAGG 180

Qy 1958 AGGAGCCCATCATCGGCGGCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCA 2017
         |||||||||||||||||
Db   344 CCGGGCGGCCAGCGCCGGCACCCTG-----AACTTCCCCAGATCACCCGTGGCAGC 397

Db 1801 AGGAGCCCATCGTGGGCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACACCGGAGACCA 1860
         |||||||||||||||||
Qy   181 CCGGCGGCCGACCGGCCAGGGCACCCGGACGGCGCCGCCAACCCCGTGGCAGC 240

Qy 2018 AGATCGGCAAGGCCGGCTACGTGACCGACCGGCGGCCGGCAGAAGATCGTGAGCCTGACCG 2077
         |||||||||||||
Db   398 GCCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCAAGGAGGCCCTGGTGGACACCGGCG 457

Db 1861 AGCTGGGCAAGGCCGGCCACGTGACCGACGGGCGCCGGCAGGTGTGAGCATCGCCG 1920
         |||||||||||||
Qy   241 GCCCCCTGGTGACCATCAGGATCGCGGCCAGCTCAAGGAGGAGCGCGTGCTGCTCGACACCGGCG 300

Qy 2078 AGACCACCAACCAGAGACCGAGCTGCAGCTCAGCTGACCCTGGCCCTGCAGGACAGCCGGCA 2137
         |||||||||||||||||||
Db   458 CCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAGCCCAAGATGATCG 517

Db 1921 ACACCCACCACCAGAGACCGAGCTGCAGGCGCCCCCTGCCACCTGGCCCTGCAGGACAGCGGCC 1980
         |||||||||||||||||||
Qy   301 CCGACGACACCGTGCTGGAGGAGATGAACCTGCCCGGCAAGTGGAGCCCAAGATGATCG 360

Qy 2138 GCGAGGTGAACATCGTGACCGAGCCCAGTACGCCGCCAGTACGCCCTCGGGCATCATCCAGGCCCAGCCCG 2197
         |||||||||||||
Db   518 GCGGGCATCGGCGGCTTCATCAAGGTGGCGCCAGTACGACCAGATCCTGATCGAGATCTGCG 577

Db 1981 TGGAGGTGAACATCGTGACCGAGCCCAGTACGCCCTCGGGCATCATCCAGCCCCAGCCCG 2040
         |||||||||||||
Qy   361 GCGGGATCGGGCGGCTTCATCAAGGTGCGGCAGTACGGACCAGATCCCGTGGAGATCTGCG 420

Qy 2198 ACAAGAGCGAGAGCCAGCTGGTGAACCAGATCATGAGCAGCTGATCAAGAGGAGAAGG 2257
         |||||||||||||||||
Db   578 GCAAGAAGGCCATCGGACCGTGCTGATCGGGCCCACCCCCGTGAACATCATCGGCCGCA 637

Db 2041 ACAAGAGCGAGAGCGAGCGTGTGAACAGCAGCCAGATCATCGAGCAGCTGATCAAGAGGAGGAAGG 2100
         |||||||||||||||||
Qy   421 GCCACAAGGCCATCGGCACCGTGCTGTGGTGGGCCCACCCCCGTGAACATCATCGGCCGCA 480

Qy 2258 TGTACCTGAGCTGGGGTGCCCGCCCACAAGGGCATCCGCAAGGTGCTGTTCCTGACAGCCAAGAGC 2317
         |||||||||||
Db   638 ACATGCTGACCCAGCTGCTGCACCCTGGAACTTCCCCATCAGCCGAGCCCTGACCGTGC 697

Db 2101 TGTACCTGGCCTGGGTGGCCCGCCCACAAGGGCGATCCGCCACAGGGCGATCCGCCACAAGGTGCTGTTCCTGACAGTGGGACAAGC 2160
         |||||||||||
Qy   481 ACCTGCTACCAGAGCCCTGACCTTCCCCATCAGCCGAGCCCTGACCGTGC 540

Qy 2318 TGGTGAGCAAGGGCATCCCGCAAGGTGCTGTTCCTGACGCGCATCGATGGCGGCATCGTGA 2377
         |||||||||||
Db   698 CCGTGAAGCTGAAGCCCGGACCCATCTGACCGGCGTGGCCCTGACCGTGC 757

Db 2161 TGGTGAGCGCCGGCATCCCCAAGGTGCTCGTTCCTGACCGTGCTGCTTCCTGAACGCGGCATCGATGGCGGCATCGTGA 2220
         |||||||||||
Qy   541 CCGTGAAGCTGAAGCCGGGATGGACCGGCGTGGCCCTGACCGTGC 600

Qy 2378 TCTACCAGTACATGGACGACCTGTACGTGGGCAGCGCCCTAGGATCGATTAAAAGC 2437
         |||||||||||
Db   758 AGAAGATCAAGGCCCTGACCGGCCATCTGCGAGGAGATGGAGAAGGGCCAAGATCACCA 817

Db 2221 TCTACCAGTACATGACGACGACCTGTACGTGGGCAGCGCCCTAGGATCGATTAAAAAGC 2280
         |||||||||||
Qy   601 AGAAGATCAAGGCCCTGACCGGCCATCTGCGAGGAGATGGAGAAGGGCCAAGATCAGCA 660

Qy 2438 TTCCCGGGGCTAGCACCGGTGAATTC 2463
         |||||||||||||||||||
Db 2281 TTCCCGGGGCTAGCACCGGTGAATTC 2306

Qy   818 AGATCGGCCCCGAGAACCCCTACAAACACCCCGTGTTGCCATCAGACAACAAGCGCACCCAGGACTTCTGGG 877
         ||||||||||||||||
Db   878 CCAAGTGCGCAAGCTGGTGAGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 937

Db   721 CCAGTGGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 780
         ||||||||||||||||
Qy   938 AGGTGCAGTGGGCATCCCCACCCCCCGCCGGCCTGAAGAAGAAGGAGCGGTGACCGTGC 997

Qy   938 AGGTGCAGTGGGCATCCCCACCCCCCGCCGGCCTGAAGAAGAAGGAGCGGTGACCGTGC 997
         ||||||||||||||||
Db   781 AGGTGCAGTGGGCATCCCCACCCCCCGCCGGCCTGAAGAAGAAGGAGCGGTGACCGTGC 840

Qy   998 TGGACGTGGGCGACGCGCACTTCAGCGTGCCCCTGACGAGGCACTTCCGCAAGTACCACCG 1057
         ||||||||||||||||
Db   841 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACAGCAAGGACTTCCGCAAGTACCACCG 900

Qy 1058 CCTTCACCATCCCCAGCATCAACAAGCACCCCGGCATCCGCTACCAGTACAACGTGC 1117
         ||||||||||||||||
Db   901 CCTTCACCATCCCCAGCATCAACAAGCACCCCGGCATCCGCTACCAGTACAACGTGC 960

Qy 1118 TGGCCCAGGGCTGAAGGCGACGCGTGCCCCTGACGTGCCCCTGACCATCTTCAGAGCAGCATGACCAAGATCCTGG 1177
         ||||||||||||||||
Db   961 TGGCCCAGGGCTGAAGGCGACGCGTGCCCCTGACGTGCCCCATCCATCTTCAGAGCAGCATGACCAAGATCCTGG 1020

Qy 1178 AGCCTTCGCGCCGCCGAACCCCGAGATCGTGATCTACCAGGCCCCCTGTACGTGGGCA 1237
         |||||||||||||||||
Db 1021 AGCCTTCCGCAAGCAGACAGAACCCGACATCGGTGATCTACCAGGCCCCCTGTACGTGGGCA 1080

Qy   170 GCGGCAAGGAGATCGGCCAGATGAAGGACTGCACCGAGGCGCCAGCCAACTTCTTCCGCG 229
         |||||||||||||||||

RESULT 5
AR373387
LOCUS   AR373387   2306 bp   DNA   linear   PAT 18-DEC-2003
DEFINITION Sequence 82 from patent US 6602705.
ACCESSION   AR373387
VERSION   AR373387.1 GI:40075490
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 2306)
  AUTHORS   Barnett,S.W., Megede,J., Greer,C. and Selby,M.
  TITLE   Expression of HIV polypeptides and production of virus-like
            particles
  JOURNAL   Patent: US 6602705-A 82 05-AUG-2003;
FEATURES   Location/Qualifiers
     source   1..2306
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match   83.1%; Score 2046; DB 6; Length 2306;
Best Local Similarity   93.6%; Pred. No. 2.4e-205;
Matches 2159; Conservative 0; Mismatches 135; Indels 12; Gaps 2;

```
Db 2378 TCTACCAGTACATGACGACCTGTAGTGGGCAGCGGCGGCCTAGGATGCGATTAAAAGC 2437
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 TCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCGGCCTAGGATCGATTAAAAGC 2280

Qy 2438 TTCCCGGGGCTAGCACCGGTGAATTC 2463
        ||||||||||||||||||||||||||
Db 2281 TTCCCGGGCGTAGCACCGGTGAATTC 2306
```

RESULT 6
AX427930

```
LOCUS       AX427930      9166 bp    DNA      linear    PAT 20-JUN-2002
DEFINITION  Sequence 168 from Patent WO0232943.
ACCESSION   AX427930
VERSION     AX427930.1    GI:21538017
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 168 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES    Location/Qualifiers
  source    1..9166
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match  82.3%;   Score 2027.6;  DB 6;  Length 9166;
Best Local Similarity  91.7%;  Pred. No. 1.4e-203;
Matches 2194;  Conservative  0;  Mismatches 179;  Indels  19;  Gaps  4;

Qy   14 TGGCCGAGCCATGAGCCACCAGGCCAACATCCTGATGCAGCGCAGCAACTTCA 73
        ||||||||| |||||||||||||||||| ||  ||       |||||||||||||
Db 2961 TGGCCGAGGCCATGAGCCACGAAGGTGAACACCAACATCATGATGCAGCAGCAACTGCA 3020

Qy   74 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGAGGGCCACATGCCCCGCA 133
        |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db 3021 AGGGCCCCAAGCGCCACCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGCCCCGCA 3080

Qy  134 ACTGCCGCGCCCCCGGCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
        ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db 3081 ACTGCCGCGGCCCCCGGCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGGTGA 3140

Qy  194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGGAAGCAGCACCCCCAAGTTCCGCC 253
        |||||||||||||||||| |||||||||||  || |||  |||||||||||||||||
Db 3141 AGGACTGCACCGAGCGCCAGGCGACAGGCTAA-TTTTTTAGGGAAGACTCTGCCTTCCCTCCCACAAGGGA 3199

Qy  254 AGGGCCCGGAGTTCCCCAGAGAACCGCGCCAACAGCCCGCAGCCGCGGAGCTGC 313
        |||||||| || | ||  ||| |||||||||  |||||||||||||||||||||
Db 3200 AGGGCCCAGGAATTTTCTTCAGAGCAGACCAGAGCAACAGCCCCCACCGGAGGCGCCAGGAGAGCTTC 3259

Qy  314 AGGTGCGCGG------------CGACAACCCCCGCCAGGAGGCCAGGGCA--- 364
        |||| ||| ||          |  |||| ||| |||||||||||||||||
Db 3260 AGGTTTGGGAAGAGACAACAACTCCCTCTCAGAGACCATAGACAAGAACTG 3319

Qy  365 ---CCCTGAACTTCCCCAGATCACCCTGTGGCAGGCCAGGGAGGAGG 421
           |  ||| || |||||||| |||||||| ||||| || |||  ||
Db 3320 TATCCTTTAGTTGCTTCCTCAGATCACTCTTTGGCAGCGACACCGGGCCGACACATAAGATAG 3379

Qy  422 GCGGGCCAGATCAAGGAGGAGGCCCTGCTGGACACCGGCAACGAGCAGATCGACAAGC 481
        || ||||||||||||||| |||||||  ||||| |||||||||| ||||  ||||||||
Db 3380 GGGGGCCAGCTGAAGGAGGAGGCCCTTCTAGACACCGGCAACGACCGGTGACAAGC 3439

Qy  482 TGAGCCTGCCCCGCAAGTGGAAGCCCAAGATGGCGGCATCGTGA 541
        |||||||||||||||||||||||||||||||||||||||||||
Db 3440 TGAACCTGCCCCGGCCGCTGGAAGCCCAAGATGATCGGCCGGCATCGTGA 3499
```

```
Db 1081 GCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGCGCACCTGCTGC 1140
        |||||||||  | ||||||| ||||||||||||||||||  || |||||||||
Qy 1298 GCTGGGGCTTCACCCACCCCCGACAAGAAGACCACAGACCAGAAGAAGG 1357

Db 1141 GCTGGGGCTTCACCACCCCCGACAAGAAGAAGAACCAGAAGAAGG 1200

Qy 1358 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGA 1417
Db 1201 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGTGCCGCCGAGAGGGACA 1260

Qy 1418 GCTGGACCGTGAACGACATCAGAGAGCTGGTGGGCAAGTGAACTGGGCCAGCCAGATCT 1477
Db 1261 GCTGGACCGTGAACGACATCCAGAGAGCTGGTGGGCAAGTGAACTGGGCCAGCCAGATCT 1320

Qy 1478 ACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCCCCAAGGCCCTGACCG 1537
Db 1321 ACGCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCCCCAAGGCCCTGACCG 1380

Qy 1538 ACATCTGCCCCCTGACCGAGGAGGAGAAGCTGGCCGAGAACCGCGAGATCCTGC 1597
Db 1381 AGGTGATCCCCCTGACCGAGGAGGAGAAGCTGGCCGAGAACCGCGAGATCCTGA 1440

Qy 1598 GCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGTGGCCGGGAATCCAGA 1657
Db 1441 AGGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGTGGGCCGGGGAATCCAGA 1500

Qy 1658 AGCAGGGCCAGGACCAGTGATGACTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGA 1717
Db 1501 AGCAGGGCCAGGGCCAGTGATGACTACCAGGATCTACCAGGAGCCCTTCAAGAACCTGAAGA 1560

Qy 1718 CCGGCAAGTACGCCCGCAAGCGCCACCGCACCGGCGCGCCACCCCCACCAACGACGACCGAGG 1777
Db 1561 CCGGCAAGTAGGCCCGCATGCCGCGCGCCACCCCCAGGAACGACGACCGAGG 1620

Qy 1778 CCGTGCAGAAGATCGCCACGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCC 1837
Db 1621 CCGGCAAGTAGGCCCGCATGCCGCGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCAAGC 1680

Qy 1838 TGCCCATCCAGAAGGAGACCTGGGACGACCTGGAGAAGGTGTGGACCGACTACTGGCACTGGA 1897
Db 1681 TGCCCATCCAGAAGGAGACCTGGGACGGCCTGGAGAAGGTACTGCGACTACTGGCACCTGGA 1740

Qy 1898 TCCCCGAGTGGGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA 1957
Db 1741 TCCCCGAGTGGGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGA 1800

Qy 1958 AGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCA 2017
Db 1801 AGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCA 1860

Qy 2018 AGATCGGCAAGGCCGGCTACGTGACCGACCGCGGGGCCCGGCAGAGATCGTGACCGACCG 2077
Db 1861 AGCTGGGCAAGGCCGGCTACGTGACCGACCGCGGGGCCCGGCAGAGGTTGGTGGCATCGCCG 1920

Qy 2078 AGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCA 2137
Db 1921 ACACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCC 1980

Qy 2138 GCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCG 2197
Db 1981 TGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCG 2040

Qy 2198 ACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG 2257
Db 2041 ACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG 2100

Qy 2258 TGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGC 2317
Db 2101 GGGGGCCAGCTGAAGGAGGAGGCCCTTCTAGACACCGGCAACGACCGGTGACAAGC 2160

Qy 2318 TGGTGAGCGCCGGGATCATCCCCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGA 2377
Db 2161 TGGTGAGCGCCGGGCATCATCCCCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGA 2220
```

```
Qy   542  TGGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCGAAGAAGGCCATCGGCCACCGTGC  601
          ||||  ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db  3500  TGGGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCCACAAGAAGGCCATCGGCCACCGTGC 3559

Qy   602  TGATCGGCGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGCTGTGACCAGTCGGCTGCA  661
          ||||  ||||||||||||||||||||||||||||||||||||||||  |||||||||||||||||
Db  3560  TGGTGGCGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGCTGTGACCAGTCGGCTGCA  3619

Qy   662  CCCTGAACTTCCCATCAGCCCCATCGCTGCGGTGCCGTGAAGCTGAAGCCCGGCATGG       721
          |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db  3620  CCCTGAACTTCCCATCAGCCCCATCGCTGCGGTGCCGTGAAGCTGAAGCCCGGCATGG       3679

Qy   722  ACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGATCAAGGCCCTGACCGCCA        781
          |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db  3680  ACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGATCAAGGCCCTGGTGGAGA        3739

Qy   782  TCTGCACCGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACA    841
          |||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db  3740  TCTGCACCGAGATGGAGAAGGAGGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACA   3799

Qy   842  ACACCCCGTGTTCGCCATCAAGAAGGACACCAAGTGGCCAAGTGGTGGACT            901
          |||||||||||||||  ||||||||||||||||||||||||||||||||||
Db  3800  ACACCCCGTGTTCGCCATCAAGAAGGACACCAAGTGGCCAAGTGGTGGACT            3859

Qy   902  TCCGGCGAGCTGAACAAGCGCACCCCAGGACTTCTGGGAGGTGCAGTTGGGCATCCCCACC    961
          ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db  3860  TCCGGCGAGCTGAACAAGCGCACCCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC    3919

Qy   962  CCGCCCGGCCGGCTTCGAAGGAGAGGAGGTGACCGTGCTGGACGTGGGCGACCGCCTACTTCA 1021
          ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3920  CCGCCCGGCCTGAACGTGACCGTGCTGGACGTGGGCGACCGCCTACTTCA            3979

Qy  1022  GGCTGCCCCTGGACGAGACTTCGGCAAGTACACGCCGCCTTCACCATCCCAGCATCAACA     1081
          |||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3980  GGCTGCCCCTGGACGAGACTTCGGCAAGTACACGCCGCCTTCACCATCCCAGCATCAACA     4039

Qy  1082  ACGAGACCCCCGGCATCGGTACCAGTACAAGTGTCTGCCCCAGGGCTGGAAGGCAGCC      1141
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db  4040  ACGAGACCCCCGGCATCGGTACCAGTACAAGTGTCTGCCCCAGGGCTGGAAGGGCAGCC      4099

Qy  1142  CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGGCCCGCAACCCCG    1201
          ||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db  4100  CCGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGGGCCCGCAACCCCG    4159

Qy  1202  AGATCGTGATCTACCA------GGCCCCCTGTACGTGGCAGCAGACCTGGAGATCGGCC    1255
          |||||||||||||||       ||||||||||||||||||||||||||||||||||||
Db  4160  ACATCGTGATCTACCAGTACATGGCCCCCTGTACGTGGCAGCAGACCTGGAGATCGGCC    4219

Qy  1256  AGCACCCGCCAAGAGTCGAGGAGATCGAGGAGCTGCGCAAGCCTGCTGCCCAGCACCTGCTCACCACCC 1315
          ||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db  4220  AGCACCCGCCAAGAGTCGAGGAGATCGAGGAGCTGCGCAAGCCTGCTGCCCAGCACCTGCTCACCACCC 4279

Qy  1316  CCGACAAGAGCACCAGAAGGAGAGGAGCCCCCTTCCTGTGATGGGATGGGTACGAGCTGCACCCCCG 1375
          ||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db  4280  CCGACAAGAGCACCAGAAGGAGAGGAGCCCCCTTCCTGTGATGGGATGGGTACGAGCTGCACCCCCG 4339

Qy  1376  ACAAGTGGACCGTGCCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGACCGTGAACGACA    1435
          |||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db  4340  ACAAGTGGACCGTGCGAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGACCGTGAACGACA    4399

Qy  1436  TCCAGAAGCTGGTGGCAAGCTGAACTGGGCCGGCCAAGGCCCGGCATCTACCCCGGCATCAAGGTGC 1495
          |||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4400  TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCCAAGGCCCCAGGCCCGGCATCTACCCCGGCATCAAGGTGC 4459

Qy  1496  GCCAGTCGTGTCCAAGCTGCTCGCGGCGCGGCCAAGGCCCGTGCCCCTGACCCG       1555
          |||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db  4460  GCCAGTCGTGTCCAAGCTGCTGCGCGGCGCGGCCAAGGCGCGAGGTGGTGCCCCTGACCCG    4519

Qy  1556  AGGAGGCCGAGCTGGAGCTGGCCGAGAATCCTGCGCGGCCCAGCCCGTGCGCGCACGGCG   1615
          ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db  4520  AGGAGGCCGAGCTGGAGCTGGCCGAGAATCCTGCGCGGCCCAGCCCGTGCGCGCACGGCG   4579
```

```
Qy  1616  TGTACTACGACCCCAGCAGCAAGGACCTGGTGGCCGAGATCCAGAAGATCCAGAGAGGGCCACGAGCAGGGCCGAGACCAGT 1675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4580  TGTACTACGACCCCAGCAGCAAGGACCTGGTGGCCGAGATCCAGAAGATCCAGAGAGGGCCACGAGCAGGGCCGAGACCAGT 4639

Qy  1676  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGA 1735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4640  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCCGCA 4699

Qy  1736  TGCGGCACCGGCCCACACCAACGACGTGGCCAGCAGCAGCAGTCGAGAAGATCGCCA     1795
          |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4700  TGAAGGGGCGCCACCAACGACGTGGCCAGCAGCAGTCGAGAAGATCGCCA           4759

Qy  1796  TGGAGAGCATCGTGATCTGGGGCAAGACCCCAAAGTTCCGCCTGCCCATCCAGAAGGAGA 1855
          ||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db  4760  CCGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCCATCCCCAGTGGGAGTTCG 4819

Qy  1856  CCTGGGAGACCTGGTGGACCGGCACTACTGCGCAGCCCACTGGATCCCCGGAGTTCG    1915
          |||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db  4820  CCTGGGAGACCTGGTGGACCGGCACTACTGCGCAGCCCACTGGATCCCCGGAGTTCG    4879

Qy  1916  TGAACACCCCCCCCGTGTGAAGCTGTGTACCAACCGCGAAGGCCGCCAAGAGCCCATCATCGGCG 1975
          ||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db  4880  TGAACACCCCCCCCGTGTGAAGCTGTGTACCAACCGCGAAGGCCGCCAAGAGCCCATCATCGGGCG 4939

Qy  1976  CCGAGACCTTCTAGTGACGGCGGCGCCGAGACCAAGATCGGCAAGGCCGGCT         2035
          |||||||||||||  |||||||||||||||||||||||||||||||||||||
Db  4940  CCGAGACCTTCTAGTGGACGGCGGCGCCGAGACCAAGCTGGCAAGGCCGGCT         4999

Qy  2036  ACGTGACCGACCGGGGGCCGCGGCCGGAAGATCGTGAGCCTGACCGACACCACCAACCAGAAGA 2095
          ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5000  ACGTGACCGACCGGGGGCCGCGGCCGGTGTGCCCCTGACCGACACCACCAACCAACCAGAAGA 5059

Qy  2096  CCGAGCTGCAGGCCATCAGCTGCCCGAGGACAGCCGCAGCGGTGAACATCGTGA      2155
          ||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db  5060  CCGAGCTGCAGGCCGCATCCAACCTGGCCCTGCGAGGACAGCAGCGGTGAACATCGTGA  5119

Qy  2156  CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC 2215
          ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db  5120  CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC 5179

Qy  2216  TGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGC 2275
          ||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db  5180  TGGTGGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCCTGGGTGC 5239

Qy  2276  CCGCCCACAAGGGCATCGGGCCGGCACCTGGAGCAGCTGGAGATCGACAAGCTGGTGAGC 2335
          |||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db  5240  CCGCCCACAAGGGCATCGGGCCGGCACCTGGAGCAGCTGGAGATCGACAAGCTGGTGAGC 5299

Qy  2336  GCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCGGCATCGTGGTGTCTACCAGTA   2387
          |||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db  5300  GCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCGGCATCGTGCAAGGCCGGAGAA 5351
```

RESULT 7
BD263706
LOCUS   BD263706   2312 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION Improved expression of HIV polypeptides and production of
            virus-like particles.
ACCESSION   BD263706
VERSION   BD263706.1 GI:33073474
KEYWORDS   JP 2002533124-A/73.
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 2312)
  AUTHORS   Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
            Greer,C., Selby,M. and Walker,C.
  TITLE   Improved expresion of HIV polypeptides and production of
            virus-like particles
  JOURNAL   Patent: JP 2002533124-A 73 08-OCT-2002;
            CHIRON CORP
            OS Artificial Sequence
            PN JP 2002533124-A/73
COMMENT

```
PD   08-OCT-2002
PF   30-DEC-1999 JP 2000591193
PR   31-DEC-1998 US      60/114495,01-DEC-1999 US    60/168471 PI
SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
KARIN HARTOG,
PI   HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
C12N5/10,
PC   C12N7/00,C12N21/02,C12N15/00,C12N5/00,A61K37/00      CC
Description of Artificial Sequence: FS(-).protmod.RTopt(+) FH  Key
                 Location/Qualifiers
FT  source           1..2312
FT                   /organism='Artificial Sequence'.
FEATURES            Location/Qualifiers
source              1..2312
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
ORIGIN
Query Match      82.2%;   Score 2025.2; DB 6;  Length 2312;
Best Local Similarity  93.3%;  Pred. No. 3.6e-203;
Matches 2156;   Conservative  0; Mismatches  138;   Indels  18;   Gaps  3;
```

[Nucleotide sequence alignment blocks (Qy/Db) follow, with position numbers. The aligned query (Qy) and database (Db) sequences are shown in paired rows with match bars (|).]

```
Db   1801 TGGAGAGGAGGCCATCGTGGGCGCCGAGACCTTCACGTGGACGGCGCCGCCACCGCG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2012 AGACCAAGATCGGCAAGGCCAAGGCCGTACGTGCTGACGTGGACGGCCGCAGAGATCGTGAGCC 2071
          |||||||||||||||  |||||||||||| |||||||||||||||||||||
Db   1861 AGACCAAGCTGGGCAAGGCCGGCGAGGCCGGCGGCGGCGGCCGGAGGTGGTGAGCA 1920

Qy   2072 TGACGCAGACCACCACCACCACCAGAGACCGAGCTGCAGGCGCCATCAGCTGACGGACA 2131
          ||||||||||||||||||||||||||||||||||||||
Db   1921 TCGCCGACACCACCACCACCACCAGAAGACCGAGCTGCAGCTGCCCTGCGACA 1980

Qy   2132 GCGGCAGCGAGGTGAACATGTGAACGCCGACGACAGCCAGTACGCCCTGGGCCATCATCCAGGCCC 2191
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GCGGCGTTGGAGCTGTGAACATGTGAACGCCGACGCCAGTACGCCCTGGGCATCATCCAGGCCC 2040

Qy   2192 AGCCCGACAAGAGCGGAGAGCGAGCTGTGAACCAGATCATCGAGACGCTGACGATCTCGCG 2251
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AGCCCGACAAGAGCGGAGAGCGAGCTGTGAACCAGATCATCGAGACGCTGACGATCTCGCG 2100

Qy   2252 AGAGAGGTGTACCTGAGCGTGGTGCCGCCCACCAGGGCATCGGCCGGCAAGCAGCAGATCG 2311
          ||||||||||||||||||||||||||||||||||||||||||||
Db   2101 AGAAAGGTGTACCTGGGCTTGGGTGCCCGGCCCACCAAGGGCATCGGCCGGCAAGCAGCAGGTGG 2160

Qy   2312 ACAAGCTGGTGACGCAAGGGCCATCGGCAAGGTGCTGTTCCTGAACGGCATCGATGCGGCA 2371
          ||||||||||||||||||  ||||||||||||||||
Db   2161 ACAAGCTGGTGACGCATCCGCAAGGGCCATCGGCATGGTGCTGTTCCTGAACGGCATCGATGCGGCA 2220

Qy   2372 TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCCGGCCCGGCCCTAGGATCGATT 2431
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCCGGCAGCCGGCCCTAGGATCGATT 2280

Qy   2432 AAAAGCTTCCCCGGGCTAGCACCGGTGAATTC 2463
          ||||||||||||||||||||||||||||||||
Db   2281 AAAAGCTTCCCCGGGCTAGCACCGGTGAATTC 2312
```

RESULT 8

AR373389

LOCUS   AR373389   2312 bp   DNA   linear   PAT   18-DEC-2003

DEFINITION   Sequence 84 from patent US 6602705.

ACCESSION   AR373389

VERSION   AR373389.1   GI:40075492

KEYWORDS

SOURCE   Unknown.

   ORGANISM   Unknown.

           Unclassified.

REFERENCE   1 (bases 1 to 2312)

   AUTHORS   Barnett,S.W., Megede,J., Greer,C. and Selby,M.

   TITLE   Expression of HIV polypeptides and production of virus-like particles

   JOURNAL   Patent: US 6602705-A 84 05-AUG-2003;

FEATURES   Location/Qualifiers

   source   1..2312

                    /organism="unknown"

                    /mol_type="genomic DNA"

ORIGIN

Query Match   82.2%;   Score 2025.2;   DB 6;   Length 2312;

Best Local Similarity   93.3%;   Pred. No. 3.6e-203;

Matches 2156;   Conservative   0;   Mismatches 138;   Indels   18;   Gaps   3;

```
Qy    170 GCGGCAAGGACGGGCCACCAGATGAAGGACTGACCGAGCGCCAGGCGCCAACTTCTTCCGCG 229
          ||||||||||||||  ||||| ||| ||||||||||||||
Db      1 GCGGCCGCGGCGAAGGACACCACCAAATGAAAGAATTCGACTGAGAACAGGCTAATTCTTCCGCC 60

Qy    230 AGGACCTGGCCTTCCTGCCGCCGCCCGGCCAAGGCCGGGCCTGGAGTTCCCCAGCGAGACCGCCA 289
          |||||||||||||||||||||||||||
Db     61 AGGACCTGGCCTTCCTGCCGCCGCCCGGGCCAAGGCCGGGCCTGGAGTTCGCCCAGCGCGACCGCCA 120

Qy    290 ACAGCCCCACCAGCCCCGCGCGGAGGTGCGCGGGC-----ACAACCCCGCCAGCGAGG 343
          ||||||||||||||||||||||||||||||||
Db    121 ACAGCCCCACCAGCCCCGCGCGCGGAGCTGCGCGGTGTGGGGCGGCGGCGAGAACAACAGCCTGAGCGAGG 180

Qy    344 CCGGCGCGGGCCGCCCAGGGCCACCCTG-----AACTTCCCCCAGATCACCCTGTGGCAGC 397
          |||||||||||||||||||||||||
```

```
LOCUS           BD263705                2300 bp    DNA       linear   PAT 17-JUL-2003
DEFINITION      Improved expression of HIV polypeptides and production of
                virus-like particles.
ACCESSION       BD263705
VERSION         BD263705.1  GI:33073473
KEYWORDS        JP 2002533124-A/72.
SOURCE          synthetic construct
  ORGANISM      synthetic construct
                artificial sequences.
REFERENCE       1  (bases 1 to 2300)
  AUTHORS       Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
                Greer,C., Selby,M. and Walker,C.
  TITLE         Improved expression of HIV polypeptides and production of
                virus-like particles
  JOURNAL       Patent: JP 2002533124-A 72 08-OCT-2002;
                CHIRON CORP
COMMENT         OS  Artificial Sequence
                PN  JP 2002533124-A/72
                PD  08-OCT-2002
                PF  30-DEC-1999 JP 2000591193
                PR  31-DEC-1998 US     60/114495,01-DEC-1999 US     60/168471 PI
                SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
                KARIN HARTOG,
                PI  HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
                C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
                C12N5/10,
                PC  C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
                Description of Artificial Sequence: FS(-).protmod.RTopt.YMWM FH
Key             Location/Qualifiers
FT   source           1..2300
FT                    /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1..2300
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match             82.0%;  Score 2019.2;  DB 6;  Length 2300;
Best Local Similarity   93.2%;  Pred. No. 1.5e-202;
Matches 2150;  Conservative    0;  Mismatches 138;  Indels   18;  Gaps    3;

Qy     170 GCGGCAAGGAGGGCCACCAGATGAGGACTGCACGGCGGCCAACTTCTTCCGCG 229
           |||||| |||||||||||||||||||||   |||||||| ||||||||||||||
Db       1 GCGGGCGGCGAAGGACACCAAATGAAAGATTGCACTGACGAGACAGGCTAATTTCTTCCGCG 60

Qy     230 AGGAACCTGGCCTTCCCCAGGGCAAGAGTTCCCCAGGCAGCAGAACCGCGCCA 289
           |||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db      61 AGGACCTGGCCTTCTGCGAGGGCAAGAGTTCAGCGAGGCGAGCAGCAGACCGCGCCA 120

Qy     290 ACAGCCCCACCAGCGCGAGCTGCAGGTGCGGGCG------ACAACCCCGAGCGAGG 343
           |||||||||||||||||||||||||||||| |||      ||||| |||||||||
Db     121 ACAGCCCCACCAGCGCCGCGCGCGAGCTGCGAGGTGTGGGGCGGCGGAGAACAACAGCCTGAGCGAGG 180

Qy     344 CCGGGGCGGAGCGCCAGGGCACCCTG------AACTTCCCCAGTACCCGTGGCAGC 397
           ||||| |||||||||||||||||||      |||||||||||||||||||||||||
Db     181 CCGGGGCGGACGCCCAGGGCACCGTGGGGCCCCCCCAGGATCACCCCCAGTACCCGTGGCAGC 240

Qy     398 GCCCCCTGGTGGCATCAAGGTGGCGGCATCGGCGGCATCGGCGGCGATGGGCGGCA 457
           |||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db     241 GCCCCCTGGTGGACCATCAAGGTGCGTGTTCCTGAACGGCATCGATGTGATGGGCGGCA 300

Qy     458 CCGACGACACCGTGCTGGGAGGAGCTGTGGAGCCCCAAGATGATCG 517
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CCGACGACACCGTGCTGGGAGGAGATGAGCCCCAAGATGATCG 360

Qy     518 GCGGCATCGGCGGCTTCATCAAGGTGCGCCTGATCGGAGATCTGCG 577
           ||||||||| |||||||||||||||||||||||||||||||||||||
Db     361 GCGGGATCGGGGGCTTCATCAAGGTTGCGGCAGTGGGGCGGCGTGGAGATCTGCG 420

Qy     578 GCAAGAAGGCCATCGGCGCCACCGGCGTGCTCAATCGGCCGCCGCA 637
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GCCACAAGGCCATCGGCGCCACCGGCGTGCTGGTGGGGCGCCCACCCCGTGAACATCATCGGCCGCCGCA 480
```

```
Db     1261 AGGACAGCTGACCGTGAACGACACATCCAGAGAGTCGTGTGGGCCAAGCTGACTGGCCAGCC 1320
Qy     1472 AGATTCACCGGGCATCAAGGTGCGCCAGCTGTGTGCAAGTGCTGCGCGGCGCCAAGGCCC 1531
Db     1321 AGATTCACGCGGGCCATCAAGGTGAGCGCAGCTGTGCAAGCTGTGCGCGGCGCACCAAGGCC 1380
Qy     1532 TGACCGACATCGTGCCCTGACCGAGGAGGGCCGAGCTGGAGCTGCCGGAGAACCGCGAGA 1591
Db     1381 TGACCGAGGTGATCCCCCTGACCCGAGGAGGGCCGAGCTGGAGCTGCCGGAGAACCGCGAGA 1440
Qy     1592 TCCTGCGCGGAGCCGTGCACGGCGTGTACTACGACACCCCAGCAAGGACCTGGTGGCCGAGA 1651
Db     1441 TCCTGAAGGAGGCGTGCACGGCGGTGTACTACGACCCCCCAGCAAGGACCTCGTGGCCGAGA 1500
Qy     1652 TCCAGAAGCAGGGCCCACGACCAGTGACCTACCAGAATCTACCAGAGAGCCCTTCAAGAACC 1711
Db     1501 TCCAGAAGCAGGGCCAGGGCCAGTGACCTACCAGAATCTACCAGAGAGCCCTTCAAGAACC 1560
Qy     1712 TGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACCACCAACGACGTGAAGCAGCTGA 1771
Db     1561 TGAAGACGGGCAAGTACGCGCCCATCGCGGCGGGCACCACCAACACCGTGAAGCAGCTGA 1620
Qy     1772 CCGAGGCCGTGCGTGCAGAGATCGCCATGGGACCATGGTGGGGCAAGACCCCCAAGT 1831
Db     1621 CCGAGGCCGTGCAGAAGGTGAGCAGCACCGCAGAGATCTGCGGCACAGTCGGGCAGATCCCCAAGT 1680
Qy     1832 TCCGCCCTGCCCATCCAGAAGGAGACCTGGGACCCTGGTGGACCTGGCCACTGGCCAGGCCA 1891
Db     1681 TCAAGTGCCCATCCAGGAGGAGAGGCCCTGGTGGTGTGATGACTGGCCAGGCCA 1740
Qy     1892 CCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCCCTGGTGAAGCTGTGGTACCAGC 1951
Db     1741 CCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCCCTGGTGAAGCTGTGGTACCAGC 1800
Qy     1952 TGGAGAAGGAGCCCATCATCGTGCGCGCCGAGACCTTCTACGTCGACGGCGCCACCCGCG 2011
Db     1801 TGGAGAAGGAGCCCATCGTGCTGGGGCGCGGAGACTTCTACGTCGACGGCGCCACCCGCG 1860
Qy     2012 AGACCAAGATCGGCAAGGCCGCTAGGTGACCGACCGGGCCGGACGGGCCGAGAAGATCGTGAGCC 2071
Db     1861 AGACCAAGTCTGGGCCAAGGCCGGCTACGTGACCGACCGGGCCGGACGGGCCGGGCAGAAGGTGGTGAGCA 1920
Qy     2072 TGACCGAGACCACCAGCAGCCAGCAGAGACCAGCTGGCCCTGCAGCTGAGAGGCCCTGCAGGACA 2131
Db     1921 TCGCCGACACCACCAGCACCAGAGACCAGCTGCCAGCTGGCCCATCCCACCTTGGGCATCATCATCCAGGACA 1980
Qy     2132 GCGGCAGCGGAGGTGAACATCGTGACCGAGCAGCAGATCATCGAGCAGATCATCGAGACCAGAAGG 2191
Db     1981 GCGGGCTGGAGGTGAACATCGTGACCAGCAGCAGATCATCGAGCAGATCATCGAGACCAGAAGG 2040
Qy     2192 AGCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGGCAACGAGCAGATCATCGGCAACGAGCAGATCATCGGCAAGAGAAGG 2251
Db     2041 AGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGGCAACGAGCAGATCATCGGCAAGAGAAGG 2100
Qy     2252 AGAAGGTGTACCTGGCGGTGGCCCGGCCCACAAGAGGGCCATCGGCAACGAGCAGATCG 2311
Db     2101 AGAAGGTGTACCTGGGGCTGGGGTGCCCCGGCCCACAAGAGGGCCATCGGCAACGAGCAGTGTGG 2160
Qy     2312 ACAAGCTGGTGATCGAGGGCAAGGGCATCCGCAAGGTGCTGTTCCTGAACGGCATGATGGGCGGCA 2371
Db     2161 ACAAGCTGGTGATCGAGGGCCAAGGGCATCCGCAAGGTGCTGTTCCTGAACGGCATGATGGGCGGCA 2220
Qy     2372 TCGTGATCTACCAGTACATGACGACTCGTATCGGGAGCAGCGGCCCCTTAGGATCGATT 2431
Db     2221 TCGTGATCTACCAGTACACGACGACTCGTATCGGGAGCAGCGGCCCCTTAGGATCGATT 2280
Qy     2432 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2463
Db     2281 AAAAGCTTCCCGGGGGCTAGCACCGGTGAATTC 2312
```

RESULT 9
BD263705

RESULT 10
AR373388

```
LOCUS       AR373388                2300 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 83 from patent US 6602705.
ACCESSION   AR373388
VERSION     AR373388.1  GI:40075491
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2300)
  AUTHORS   Barnett,S.W., Megede,J., Greer,C. and Selby,M.
  TITLE     Expression of HIV polypeptides and production of virus-like
            particles
  JOURNAL   Patent: US 6602705-A 83 05-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..2300
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

(Nucleotide sequence alignment, Query (Qy) vs Database (Db), positions 638–2300)

```
Query Match       82.0%;  Score 2019.2;  DB 6;  Length 2300;
Best Local Similarity  93.2%;  Pred. No. 1.5e-202;
Matches 2150;  Conservative  0;  Mismatches 138;  Indels  18;  Gaps  3;

Qy   170 GCGGCAAGGAGGGCCACCAGATGAAGGAGGACTGACCGGCGCCAGGCGCCAACTTCTTCCGGCG 229
         ||||   |||  |||    |||||| |||   | ||||| |  | | |||||||| |||||||||
Db     1 GCGCGGCCGCCGAAGGACACCCAAATGAAAGATTGCACTGAGAGACAGGCTAATTCTTCCGGCG 60

Qy   230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCAGCGAGCAGAACCGGCCCA 289
         |||||||||||||| |  |||||||||||||||||||| |||||||||||||||| || ||
Db    61 AGGACCTGGCCTTCCTGCAGGGCAAGGCCCGCGAGTTCAGCGAGCAGAACCCGCCCA 120

Qy   290 ACAGCCCACCAGCGCGGAGCTGCAGGTGCGCGCG-----ACAACCCCGCAGCGGAGG 343
         |||||||||  ||  | ||||||||||||| || |     ||||| | | |||| |||
Db   121 CAAGCTGCACCCCGCCGCGGAGCTGCAGGTGTGGGGCGGCGGAGAACAACAGCCTGAGCCGGAG 180

Qy   344 CCGGCGCCGAGCGCCAGGCACCCTG-----AACTTCCCCCAGATCACCCTGTGCCAGC 397
         |||| |||||||| |||| |||| |      |||||||||||||||||||||| ||||
Db   181 CCGGCGCCGCCGGAGCACCGTGAGCGTCACCAACTTCCCCCAGATCACCCTGTGCCAGC 240

Qy   398 GCCCCTGTGAGACACCGGTCTGGGCGCCAGATCAAGGAGCCCTGCTGACACCGGCG 457
         ||||||||| ||| ||| || ||||  ||| ||||||||||||||| | |||||||
Db   241 GCCCCTGTGACCACCGGTCTGGGCGACATCGGGCGCCTCAAGGAGCCCTGCTGCACACCGGGCG 300

Qy   458 CCGACGACACCGGTGCTGAGGAGATGAGCCTGCCCGGCAAGTGGGAGGCCCAAGATGATCG 517
         ||||||||||||| |||||||||||||  |||| || ||||||||||||||||||||||||
Db   301 CCGACGACACCGGTGCTGAGGAGATGAACCTGCCCGGCAAGTGGGAAGCCCAAGATGATCG 360

Qy   518 GCGGCGCGCTTCATCAAGGTGCGCCAGTAGCCTGGCCGCCAGATCCTGATCGAGACTGCGG 577
         ||||| ||||||||||||||||| ||| |||||||||  |||||||| ||||||||||||
Db   361 GCGGATCGGGCTTCATCAAGGTGCCGCAGTAGCCTGCGCCAGATCCCGTGAGATCTGGCG 420

Qy   578 GCAAGAAGCCATCGGACCGTGCTGATCGGCCCCACCCCCCGTGAACATCATCGGCCGCCA 637
         |||||||||||||||||||||||||||||||||||| |||| ||||||||||||||||||
Db   421 GCCACAAGCCATCGGACCGTGCTGGGATGACCGAGTTCAAGGCTCAGGAGCATCATCGGCCGCA 480

Qy   638 ACATGCTACCCAGCTGGGCTGGCTGCCACCCTGACACCCTGAACTTCCCATCAGCGCCGTGC 697
         ||||||| |||||||||||||  ||||||||||||||||| |  | |||||||| |||||||
Db   481 ACCTGCTACCCAGCTGGGCTGCGCCCTGGTGTGACCCTGAGACTTCCCATCAGCGCCGGTGC 540

Qy   698 CCGTGAAGCTGAAGCCCGGATGGCGGATGCCGGGATGACGGCAGTGCGCCCTGACCGAGG 757
         ||||||||||||||||  |||||||||||||  |||| |||||||||||||||||||||||
Db   541 CCGCAAGCTGAAGCCCGGATGGCGGATGAGCCAGAGCAGTGCGCCCTGACCGAGG 600

Qy   758 AGAAGATCAAGGCCCTGACCGGCCATCTGCGAGGAGTGAGAAGGAGGGGCAAGATCACCA 817
         |||||||||||||||| |  |||||||| ||||||||| |||| ||||||||||||||||
Db   601 AGAAGATCAAGGGCCCTGGTGAGAGATCTGCACCGGAGATGGAGGGCGGAGATCAGCA 660

Qy   818 AGATCGGCCCCGAGAACCCTACAACACCCCCGTGTTCGCCATCAAGCGCACACGAGGACAGCA 877
         ||||| |||||||||||  ||||||||||| ||||||||||||||||||||||||||||||||
Db   661 AGATCGGCGCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGCGCACAACGAGGACAGCA 720

Qy   878 CCAAGTGGCAGCTGCAAGCTTCCGCGACTTCGGCGCGCCATCGAACAAGCGCACCCCAGGACTTCTGGG 937
         |||||||||||||||||||| |  |||||||| ||||| |||||||||||||||||||| ||| |||||
Db   721 CCAAGTGGCGCAGCTGCAAGCTCCGCGACTTCGGCGCCATCGAACAAGCGCACCCCAGGACTTCTTGGG 780

Qy   938 AGGTGCAGTGGGCATCCCCACCCCCCGGCCTGAAGAGAAGAGAGGGGTGACCGTGT 997
         ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   781 AGGTGCAGTGGGCATCCCCACCCCCCGGCCTGAAGAGAAGAGAGGAGCCGTGACCGTGT 840

Qy   998 TGGACGTGGGCGACGCCTACTTCAGCGGTGCCCCTGACGAGGGACTTCCGCAAGTACACCG 1057
         |||||||||||||||||||||||||||||| |||||||||||||||| ||||||||||||
Db   841 TGGACGTGGGCGACGCCTACTTCAGCGGTGCCCCTGACGAGGGACTTCCGCAAGTACACCG 900

Qy  1058 CCTTCACCATCCCCAGCATCAACAACAACGACATCAACAACAACGACACCCGGCATCCGCCTACCAGTACAACGTGTC 1117
         |||||||||||||||||||||||| ||||||||||  |||||||||||||||||||||||||||||||||||||
Db   901 CCTTCACCATCCCCAGCATCAACAACAACGACATCAACAACAACGACACCCGGCATCCGCCTACCAGTACAACGTGTC 960

Qy  1118 TGCCCAAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAACAAGATCCTGG 1177
         ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAACAAGATCCCTG 1020

Qy  1178 AGCCCTTCGCGGCCGCCGCAACCCGAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCA 1237
         ||||||||| | ||||||| ||| | ||||||||| ||||||||||||| |||||||| ||
Db  1021 AGCCCTTCGGCAAGCAGCAGATCGTGATCGTGATCTACCACCAGGCCCCCGTACGTGGGGGCA 1080

Qy  1238 GCGACCTGGAGATCGGCCAGCAGCCGCCAAGATCGAGAGCTGCCAAGCACGCTGCTGC 1297
         |||||||||||||||||||||||| |||||||||||||||||| ||| ||| ||||||
Db  1081 GCGACCTGGAGATCGGCCAGCAGCCGCCACCAAGATCGAGAGCTGCCACCAGCACCTGCTGC 1140

Qy  1298 GCTGGGGTTCACCACCCCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGTGGATGG 1357
         |||| |||||||||||||||||||||||||||||||||||| ||| |||  |    |
Db  1141 GCTGGGGGTTCACCACCCCGACAAGAAGCACCAGAGGAGCCCCCTTCCTGCCCAT-- 1198

Qy  1358 GCTACGAGCTGACACCCGACAGTGACCGTGCAGCCATCGAGCTGCCCGAGAAGGAGA 1417
         ||| |||||||||  |||||| |||||||| |  ||||||||||  |||||| ||||
Db  1199 ----CGAGCTGACCACCCGACAAGTGACCGTGCAGCCATCATGCTGCCCGAGAAGGACA 1254

Qy  1418 GCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGAAACTGGGCCAGCCAGATCT 1477
         ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  1255 GCTGGACCGTGAACGACATCCAGAGAGCTGGTGGCAAGCTGAAACTGGGCCAGCCAGATCT 1314

Qy  1478 ACCCGGCATCAAGGTGCGCCAGTCTGCTGCGCGGCCGCCAAGGCCCTGACCG 1537
         |||||||||||||| ||| ||||||||||||| ||| ||||||||||||||||
Db  1315 ACGCCGGCATCAAGGTGAAGCAGTCGTGCAAGTGCTGCGCGGCACCAGGCCCTGACCG 1374

Qy  1538 ACATCGTGCCCCTGACCGGCGTGACTACGACCCCAGCAAGTGGTGGCCGAGATCCTGC 1597
         |||| ||||||||||| |||| |||||||||||||||||||||| ||||||||||||||
Db  1375 AGTGATCCCCCTGACCGGCGAGAGTGTACTACCAGCAAGCAAGGCCTGGAGCTGAGATCCTGA 1434

Qy  1598 GCGAGCCCGTGACAGGCGTGTACTACGACCCCAGCAAGAAGAACCTGTGTGGCCGAGATCCAGA 1657
         |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  1435 AGGAGCCCGTGCGCGAGGTGTACTACCACCCCAGCAAGAACCTGGTGGTCCGAGATCCAGA 1494

Qy  1658 AGCAGGGCCACGACCACCAGTGGACCTACAGATCTACCAGGATCTACCAGAACAACCTGAAGA 1717
         ||||||||||||||||| |||||||||| ||||| ||||| |||||||||||||||||||||
Db  1495 AGCAGGGCCACGACCACCAGTGGACCTACCCAGATCTACCAGGATCTACCAGGAGCCCTTCAAGAACAACCTGAAGA 1554

Qy  1718 CCGGCAAGTAGCTGACCCTGGCGTGATCGGCCATGCCCATCAGCGCCCAAGTTCCGCC 1777
         |||||||||||| ||| ||||||||||  |||||||||  || |||||||||| |||||
Db  1555 CCGGCAAGTAGCCTGCGGATGAACTCTTCAAGCTCAGGAGCGCGCCCAAGTCAAGC 1614

Qy  1778 TGCCCATCCAGAAGATCGACCTGGCGGGCAATGGCCGCCCATCTGGTGACCACTACCCTGGGAA 1837
         ||||||||| ||||||| |||| |||||| ||| ||||| ||| ||||| ||| ||||| ||||
Db  1615 TGCCCATCCAGAAGATCGGCACCGACGGGCAGTGGCCGCCCATCTGGTGACCACTACCAGTGGGAA 1674

Qy  1838 TCCCCGAGTGGGAGTTCGTGAACACCCCCCCGTGTTCTACGTGGAGCTGTGTGGTACCAGTGGGAGA 1897
         ||||||||||||||||||||||||||| ||||| ||| |||||||||||| ||||||||||||||||
Db  1675 TCCCCGAGTGGGAGTTCGTGAACACCCCCCCGTGTTCTACGTGGAGCTGTGGTACCAGTGGGAGA 1734

Qy  1898 AGGAGCCCATCATCGGCGCCGGACTTCCGCGCGCCCGAGAACAAGCGGCACACGAGGAGGACCA 1957
         |||||||||||||||| ||||||| ||| ||| ||||||||||||||| |||||||||||||||
Db  1735 AGGAGCCCATCATCGGCGGCCGGACTTCCGCGCGCCCGAGAACAAGCGGGCCGCCACCGAGACCA 1794

Qy  1958 AGATCGGCAAGGCCGGCTACGTGACCGGCGGCGAAAGATCGTGAGCGTGCCCTGACCG 2017
         ||||| |||||||||||||||||||||| |||||||| |||||||||||||||||||||
Db  1795 AGTCGGGCAAGGCCGGCTACGTGACCGGCGGCGCCGCAGAAGAGTGGTGAGCGTGCCCACCA 1854

Qy  2018 AGATCGGCAAGGGCGGCATCCCCACCCCACCCCGGCCGGAGCTGAGGGCGGAGAAGATCGTGAGCCGGA 2077
         ||||| ||||| |||||||||||||||| || ||||||||||||| |||||||| | ||| ||||||||
Db  1855 AGTCGGGCAAGGGCGGCATCCCCACCCCACCCCGGCCGGAGCTGAGGGCGGAGAAGAGTGGTGAGCATCGGCCC 1914

Qy  2078 AGACCACCAACCAGAGACCGAGCTGCGCAGCACCGGCGGCA 2137
         || || ||||| |||||||||||||||| ||||||||||||
Db  1915 ACACCACCACCAACCAGAGACCGAGCTGCACCTGGCCCTGCAGAACAGCGGCC 1974

Qy  2138 GCGAGGTGAACATCGTGACCAGCAGCACACGTGACCCTGCCGGCATCATCCAGGCCCCCAGCCCG 2197
         |||| ||||||||| |||| ||||||||||| ||||||||| ||||||||| |||||||||||||
Db  1975 TGGAGGTGAACATCGTGAACCAGCAGCACAGCC ||||||||||||||||||| 2034

Qy  2198 ACAAGAGCGAGAGCAGCGAGCTGGTGAACCAGACCAGCATGGTGAACAGTCGTGATCAAGAAGAGAAAG 2257
         ||||||||||| ||||||||||  ||||||||||||  ||  ||||| ||||||| ||||||||||||
Db  2035 ACAAGAGCGAGAGCAGCGAGCTGGTGAGCTGAGCATCATCGAGCTGATCATCGAGCAGAAGGAGGAGAAGG 2094

Qy  2258 TGTACCTGAGCTGGGTGCCCGCCCAAGGGCATCGGCGGCAACGAGCGATCGACAAGC 2317
```

```
Db  2095  TGTACCTGGCTGGGTGCCGGCCCGCCCACAAGGGCATCGGCCGGCATCGGCCAACGAGCAGGTGACAAGC  2154
                                    ||||||||||||||||||||||||||||||||||||||||||||
Qy  2318  TGGTGAGCCAAGGGCATCGGCCATCGGCCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGA   2377
          |||||||||||||||||||||||
Db  2155  TGGTGAGCGGCCGGCATCCGCCAAGGTGCTGTTCCTGAACGGCATCGATGGCGGCATCGTGA          2214

Qy  2378  TCTACCAGTAGTACATGAGCAGACCTGCAGCGGCGGCGGCCGGCGCGGCCCTAGGATCGATTAAAAGC    2437
          |||||||||  ||||||||||||
Db  2215  TCTACCAGTACATGAGCAGACCTGCAGCGGCGGCCGGCGCGGCCCTAGGATCGATTAAAAGC         2274

Qy  2438  TTCCCGGGCTAGCACCGGTGAATTC  2463
          |||||||||||||||||||||||||
Db  2275  TTCCCGGGCTAGCACCGGTGAATTC  2300
```

RESULT 11
AX427936

LOCUS   AX427936   9788 bp   DNA   linear   PAT 20-JUN-2002
DEFINITION Sequence 174 from Patent WO0232943.
ACCESSION   AX427936
VERSION   AX427936.1 GI:21538023
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE   Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 174 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES   Location/Qualifiers
     source   1..9788
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match   82.0%; Score 2019.2; DB 6; Length 9788;
Best Local Similarity 91.6%; Pred. No. 1e-202;
Matches 2190; Conservative 0; Mismatches 178; Indels 24; Gaps 4;

```
Qy    14  TGGCGGAGCCATGAGCCAGGCCAACATCCGCTACCAGCAGCAACTTCA                       73
          |||||||||||||||||||||||||||| ||| ||    |||  ||||
Db  2961  TGGCCGGAGCCATGAGCCAGGCCAACATCATGAGCCAGCAACTGCAACAACCAACATCATGAGCCAGCAACTGCA  3020

Qy    74  AGGGCCCCAAGGCCGCCCCCCCCCGAGAAGTGCCACCAGATGA                           133
          ||||||||||||||||| |||   |||||||||||||||||||
Db  3021  AGGGCCCCAAGGCCGCCACCATCAACTGCTTCAACTGCTTCAACTGGCTGCAAGAGGGCCACCAGATGA  3080

Qy   134  ACTGCCGCGGAGGCGCCCCCCCCCCAAGAAGGCGGCAAGAGGAGGCAACTGCACCGGA            193
          ||||||||||||||||            ||||||||||| ||||||||||||
Db  3081  ACTGCCGCGGAGGCGCCCAAGAAGGCGCGGCAAGAGGAGGGCAAGAGGAGGCCACCAGTGA        3140

Qy   194  AGGACTGCACCGGACCGCCAACTCTTCCTCCGCGGAGAAGTGCCTTCCCGAGGACCTGGCCTT       253
          ||||||||||||    ||||||||||||||
Db  3141  AGGACTGCACCGGCCGCCTAA------TAGGGGAAGATCTGGCCTTCCCACCAGGGGA          3194

Qy   254  AGGGCCGCGGAGTTCCCAGCGAGGACAACCGCCGAGAACCGCGAGCTGC                      313
          |||||| || ||||   ||||||||| |||||||||||||||||||
Db  3195  AGGGCCAGGGAATTTTCTTCAGACCAGACAGCAACAGCCCCACCAGAGAGCTTC                3254

Qy   314  AGGTCGCGG------CGACAACCCCCGCAGCGAGCGCCGGCCGCAGCGGGCA---              364
          ||||  || |||   ||||||||||||||||||
Db  3255  AGGTTTGGAAGAGAACAACTCCCTCTTCAGAGAGCAGGAGCAGGAGCAGGAGGACTG            3314

Qy   365  ---CCCTGAACTTCCCTCAGATCACCCGTGGCAGCATCCGGCATCAAGGTGG                   421
             ||||| ||||||||||||||||||||||||||||||||||||||||||
Db  3315  TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCACCCCTCGTCACAATAAAGATAG          3374

Qy   422  GCGGGCCAGATCAAGGAGAGGCCGCCCGCCGGCACCGGTGCTGGACCACC                     481
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3375  GGGGCCAGTGAGTCGACGAGGAGCCCTTCTAGACACCGGCGCCGACACCGTGCTGTGGAGAGA        3434

Qy   482  TGAGCCTGCCCGGCCAAGTGAAGCCAACGACGATGATCGGCGGCATCGGCGGCTTCATCAAGG      541
          |||  |||||||||||||||||||||||||     |||||||||||||||||||||||||||
Db  3435  TGAACCTGCCCGGCCAAGTGAAGCCAACGACGATGATCGGCGGCATCGGCGGCTTCATCAAGG      3494

Qy   542  TGCGGCAGTACGACCAGATCCTGATCGAGATTCGGCGGCGCCACAAGAAGGGCCATCGGCACCGTGC   601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3495  TGGGGCAGTACGACCAGATCCTGATCGAGATCTGGCGGCGCCACAAGAAGGGCCATCGGCACCGTGC   3554

Qy   602  TGATCGGCCCCACCCCGTGAAACATCATCGGCGCCAACATGTGACCCAGCTGGGCTGCA          661
          |||||||||||||||||     |||| ||||| |||||||||||||||||||
Db  3555  TGGTGGGCCCCACCCCGTGAACATCATCGGCGCCAACATCGGCTGACCCAGCTGGGCTGCA        3614

Qy   662  CCCTGAACTTCCCCATCAGCCCATCGACCCGTGCCCGTGAAGCTGAAGCCCGGCCATGG          721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3615  CCCTGAACTTCCCCATCAGCCCATCGACCCGTGCCCGTGAAGCTGAAGCCCGGCCATGG          3674

Qy   722  ACGGGCCCAAGGTGAAGCAGTGCCCCTGACCGAGGAGATCAAGGCCCTGACCGCCA            781
          |||||||||||||||||||||| |||||||||||||||||||||| |||| |
Db  3675  ACGGGCCCAAGGTGAAGCAGTGCCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGA         3734

Qy   782  TCTTGCGAGGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCAGAGACCCCTACA          841
             | ||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db  3735  TCTGCACCGGATGGAGAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACA          3794

Qy   842  ACACCCCGTGTTCGCCATCAAGAAGAAGAAGAGCACCAAGTGGCGCAAGTGGTGGACT          901
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3795  ACACCCCGTGTTCGCCATCAAGAAGAAGAAGAGCACCAAGTGGCGCAAGTGGTGGACT          3854

Qy   902  TCCGCGGAGTGAACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGCATCCCCACC          961
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3855  TCCGCGGAGTGAACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGCATCCCCACC          3914

Qy   962  CCGCCGGCCTGAAGAAGAGAGAAGAGCGTGACCGTGCGTGCCGACGTGGGCGACGCCTACTTCA     1021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3915  CCGCCGGCCTGAAGAAGAGAGAAGAGCGTGACCGTGCGTGCCGACGTGGGCGACGCCTACTTCA     3974

Qy  1022  GCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACA          1081
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3975  GCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACA          4034

Qy  1082  ACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCC       1141
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  4035  ACGAGACCCCCGGCATCCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCC      4094

Qy  1142  CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCCCCGCCAACCCCG      1201
          ||||||||||||||  |||||||||||||||||||||  |||||||  ||||||
Db  4095  CCGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGGAAGCAGAACCCCG       4154

Qy  1202  AGATCGTGATCTACCA------GGCCCCCCGTGTACGTGGGCAGCGAGATCGGCGCC         1255
          | |||||||||||||      ||||||||||  |||||||| |||||||||||||||
Db  4155  ACATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCC       4214

Qy  1256  AGCACCGCCAAGATCGAGGAGCTGCGCGGCCACCTGCTGCGCTGGGGCTTCACCACCC        1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4215  AGCACCGCCAAGATCGAGGAGCTGCGCGGCCACCTGCTGCGCTGGGGCTTCACCACCC        4274

Qy  1316  CCGACAAGAAGCACCAGAAGAGGAGCCCCTTCCTGTGATGGGCTACGAGCTGCACCCCG       1375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4275  CCGACAAGAAGCACCAGAAGAGGAGCCCCTTCCTGTGATGGGCTACGAGCTGCACCCCG       4334

Qy  1376  ACAAGTGGACCGTGCAGCCCATCGTGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACA      1435
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  4335  ACAAGTGGACCGTGCAGCCCATCCGCTGCCCGAGAACGACAGCAGCCTGGACCGTGAACGACA    4394

Qy  1436  TCCGAAGCTGGTGGCAAGCTGGCCAGATCTACCCGGCATCAAGGTGG                     1495
          ||| |||||||||||||||||||||||||||||| ||||||||||||
Db  4395  TCCAGAAGCTGGTGGCAAGCTGGCCAGATCTACCCGGCATCAAGGTGC                    4454

Qy  1496  GCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCCG       1555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   4455 GCCAGCTGTGCAAGCTGCTGCGCGGGCACCCAAGGCCCTGACCGAGGTGTGCCCCTGACCCG 4514
                 ||||||||||||||||||||| |||||||||||||
Qy   1556 AGGAGGCCGAGCTCGAGCTGGCCGAGATCCTGCGCGGAGCCCGTGCACGGCG 1615
                 |||||||| |||||||||||||||||||||||
Db   4515 AGGAGGCCGAGCTCGAGCTGGCCGAGATCCTGCGCGGAGCCCGTGCACGGCG 4574

Qy   1616 TGTACTACGACCCCAGCAAGGACCTGGTTGGCCGAGATCCAGGAGCAGGGGCCACGACCAGT 1675
                 ||||| |||||||||||||||||||| |||||||||||||||||||||||||
Db   4575 TGTACTACGACCCCAGCAAGGACCTGATCGCCGAGATCCAGGAGCAGGGGCCACGGGCCAGT 4634

Qy   1676 GGACCTACCAGATCCACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGGCCAAGA 1735
                 |||||||||||| |||||||||| ||||||||||||||| |||||||||||||||||
Db   4635 GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGGCCCCGCA 4694

Qy   1736 TGCGCACCGCCCACCAACCAACGAGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCA 1795
                 |||||| |||||||||||||||| |||||||||||||||||||| |||||||||||||
Db   4695 TGAAGGGCCGCCACCAACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCA 4754

Qy   1796 TGGGAGAGCATCGTCGATCTGGGCAAGACACCCCAAGTTCCGCCTGCCATCCAGAAGGAGA 1855
                 ||||||||||||||||||||||||||||| ||||||||| |||||||||||||||||||
Db   4755 CCGAGAGCATCGTCGATCTGGGCAAGACCCCCCAAGTTCAAGCTGCCATCCAGAAGGAGA 4814

Qy   1856 CCTGGGAGACCTGGTGACCGACTACTGGCAGCCACCTCCCGAGTGGGAGTTCG 1915
                 |||||||||||||||||||||||||| ||||||| ||||||||||||||||||
Db   4815 CCTGGGAGGCCTGTGGTGGACCGAGTACTGGCAGCCACCTCCCGAGTGGGAGTTCG 4874

Qy   1916 TGAACACCCCCCCCCTGTTGTGAAGCTGTGTACCAGTGGAAGAAGGAGCCCATCATCGGCG 1975
                 ||||||||||| ||||||||||||||| ||||||||||||||||| |||||||||||
Db   4875 TGAACACCCCCCCCCTGTTGTGAAGCTGTGTACCAGTGGAAGAAGGAGCCCATCATCGGCG 4934

Qy   1976 CCGAGACCTTCTACGTGACGTGGACGGGGCCGCCAACGCAGATCGGCAAGGCGCGGCT 2035
                 |||||||||||||||||||| ||||||||||||||||| |||||||||||||||||||
Db   4935 CCGAGACCTTCTACGTGACGAGCGGGCCGCCGCCAACGCAGATCTGGGCAAGGCCGGCT 4994

Qy   2036 ACGTGACCGACCGGGGCCGGCCCAGAAGATCGTGACGCCTGACCGAGACCACCAGAAGA 2095
                 |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   4995 ACGTGACCGACCGGGGCCGGCCCAGAAGATCGTGACGCCTGACCGAGACCACCAGAAGA 5054

Qy   2096 CCGAGCTGCAGGCCATCCAGCTCGCCCGCAGACAGCGGCGCAGCGGAGCTGAACATCGTGA 2155
                 ||||||||||||| ||||||||||||||||||||||||||| |||||||||||||||||
Db   5055 CCGAGCTGCAGGCCATCCAGCTCGCCCGCAGACAGCGGCGCAGCGGAGCTGAACATCGTGA 5114

Qy   2156 CCGACAGCCAGTACGCCCTGGGCATCTCAGGCCCAGCCCGACAAGAGGCGAGAGGCGAGC 2215
                 ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   5115 CCGACAGCCAGTACGCCCTGGGCATCCTCAGGCCCGACAAGAGGCGAGAGGCGAGC 5174

Qy   2216 TGGTGAACCAGATCATCGAGCTGATCATGAAGAAGGAAGTGTACCTGAGCTGGGTGC 2275
                 ||||||||||||||||||||||||||||||| ||||| ||||||||||||| ||||||
Db   5175 TGGTGAGCCAGATCATCGAGCTGATCATGAAGAAGGAAGTGTACCTGCCTGGTGC 5234

Qy   2276 CCGCCCACAAGGGCATCGGCGGCGCCAACCAGCAGATCGACACAAGCTGGTGACCAAGGCCATCC 2335
                 |||||||||| |||||||||||||||||||||||||||| ||||||||||||||||
Db   5235 CCGCCCACAAGGGGCATCGGCGGCGCCAACCAGCAGTGACCCTGGTGAGCGCCGCCCATCC 5294

Qy   2336 GCAAGGTGCTGTTCCTGGACGACGCCATCGATGGCGGCATCGTGATCTGATCTACCAGTA 2387
                 |||||||||||| |||||||||||||||||| ||| |||||
Db   5295 GCAAGGTGCTGTTCCTGGACGACGCCATCGATGGCGGCATCGTGATCTGATCTACCAGTA 5346
```

Query Match   81.4%;   Score 2005.2;   DB 6;   Length 9169;
Best Local Similarity   91.4%;
Matches 2187;   Conservative   0;   Mismatches 183;   Indels 22;   Gaps 5;
                 Pred. No. 3.1e-201;

```
Qy    14 TGGCCGAGGCCATGAGCCAGGCCACAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA 73
             ||||||||||||||||| |||||||||||||||| |||||||||| |||||||||||
Db  2967 TGGCCGAGGCCATGAGCCAGGCCACCAGCACACAACAATCATGATGCAGCGCGGCCAACTTCA 3026

Qy    74 AGGGCCCCAAGGCCATCATCAAGTGTTCAACTGCGCGCAAGGAGGGCCCATCGCCCGCA 133
             ||||||||||||||||| |||| |||| |  ||| |||||||||| ||||||||||||
Db  3027 AGGGCCAGAAGGCGCAT---CAAGTGCTTCAACTGCGGCAAGGGCCCACCTGGCCCGCA 3083

Qy   134 ACTGCCGGCCCCCCCCCGCGCAAGAGAGGGCTGCTGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
             |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3084 ACTGCCGGCCCCCTGCGCGCAAGAGAGGGCGTGCTGGAAGTGGGCGGCAAGGAGGGCCACCAGATGA 3143

Qy   194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGAGACCTGGCCTTCCCCCAGGGCA 253
             |||||||||| ||||||||| |||||||||||| ||||||||||||||||||||||||||
Db  3144 AGGACTGCACCGAGCGACCAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCCCACAGGGA 3202

Qy   254 AGGCCCGCGAGTTCCCAGCAGCAGAACCCGCCCAACAGCCCACCAGCCGCGAGCTGC 313
             ||||||||| |||||| |||||| ||| |||||||| || |||||||||||||||||
Db  3203 AGGCCAGGAATTTCTTCAGCAGCAGACCAGCAGAGCCAACAGCCCCACCAGCAAGGAGAGCTTC 3262

Qy   314 AGTGCGCGG------CGACAACCCCGCAGCGAGGCGGGCCGCCAGGGCA-- 364
             ||||||||       ||||||| ||||| ||||||||||||||||||
Db  3263 AGGTTTGCGGGAAGAGACAACAACTCCCTCGAGAAGCAGGCAGGAACAAGGAACTG 3322

Qy   365 ---CCCTGAACTTCCCCCAGATCACCCGTGTGCAGCGCAGGTGCAGGCATCAAGGTGG 421
                ||||||||||||||| |||||||||| |||||||||||| |||||| |||||||
Db  3323 TATCCTTTAGCTTCCCTCGAGTCACTCTTTGCAGCGACCCTCGTCACAATAAGATAG 3382

Qy   422 GCGGCCAGATCAAGGACGAGGCCCTGCTGCGACACCGTGCTGAGGAGA 481
             ||||||||||||||||||||||||||||||||||| ||| |||||||||||
Db  3383 GGGGCCAGCTGAAGGAGGCCCCTTCAGACACCGCCCGACGACACCGTGGGGAGGA 3442

Qy   482 TGAGCCTGCCCGGCAACAGTGGAAGCCCAAGAGATCGAGATCTGGTGC 541
             ||||||| |||||||||||||||||||| ||||||||||| ||||||||||
Db  3443 TGAACCTGCCCGGCCGCCAAGATGATCTGAGATCATGCGGCGGCTTCATCAAG 3502

Qy   542 TCGCCAGTACGACCAGATCCTGATGCAGCCATCATCGGCGCCCAACATGTCTGAACCGTGC 601
             ||||||||||||||||||||| |||||||||||||||| |||||||||||||||||||
Db  3503 TGGGCCAGTACGACCAGATCCTGATGCAGCCACAAGGCCATCCAGAAGGCCATCGCCACCGTGC 3562

Qy   602 TGATCGGCCCCATCAGCCCCATCCGAACTGCTGGGCTGGGCTGCA 661
             |||||| ||||||||||||||||| |||||||| ||||||||||
Db  3563 TGGTGGGCCCACCCCGTCATCCGAACATCATCGCTGACCCAGATCGGCTGCA 3622

Qy   662 CCCTGAACTTCCCCATCAGCCCCATCAGCCCCATCGAGAGCCCGTGAAGCCCGGCATGG 721
             ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  3623 CCCTGAACTTCCCCATCAGCCCCATCAGCCCCATGCCCGTGAAGCTGAAGCCCGGCATG 3682

Qy   722 ACGGCCCCAAGGTGAAGCAGTGCCCCTGACCGAGAAGATCAAGGCCCTGACCGGCCA 781
             ||||||||||||||||||||| |||||||||||||||||||||||| |||||||||
Db  3683 ACGGCCCCAAGGTGAAGCAGTGGCCGTGCCCCTGACCGAGGAGAATCAAGGCCCTGTGGAGA 3742

Qy   782 TCTGCGAGGATGGAGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACA 841
             | |||||||||||| |||||||||||||||||||||||| |||||||||||||
Db  3743 TCTGCACCGAGGATGGAGGAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACA 3802

Qy   842 ACACCCCCGTGTTCGCCATCAAGAAGCAGCAGAGCCAAGTGCGGCAAGCTGTGGACT 901
             ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  3803 ACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCAGCACCAAGTGGCGCAAGCTGGTGGACT 3862
```

RESULT 12
AX427931

```
Qy   1976  CCGAGACCTTCTACGTGGACGGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCGGCT  2035
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4943  CCGAGACCTTCTACGTGGACGGCCGCGCCCCAACCGCGAGACCAAGATCTGGCAAGGCCGGCT  5002

Qy   2036  ACGTGACCGACCGGCGGCCGGCCGCCAGAAGATCGTGACCGAGACCGACGACCACCAGAAGA  2095
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5003  ACGTGACCGACCGGCGGCCGCGCCCAGAAGTGGTCCCCTGACCGACCACCACCAGAAGA  5062

Qy   2096  CCGAGTCGCAGGCCATCCAGCTGGCCTGGCCTGCGAGGACGCGGCGAGGTGAACATCGTGA  2155
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5063  CCGAGTCGCAGGCCATCCAGCTGGCCTGGCCTGCGAGGACAAGCGGCCTGGAGGTGAACATCGTGA  5122

Qy   2156  CCGACAGCCAGTACGCCCCTGGGCATCATCCAGCCCCAGCCCGACAAGAGCGAGAGCGGAGC  2215
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5123  CCGACAGCCAGTACGCCCCTGGGCATCATCCAGCCCCAGCCCGACAAGAGCGAGAGCGGAGC  5182

Qy   2216  TGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGTGTACCTGACCGAGCTGGGTGC  2275
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5183  TGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGTGTACCTGGCCTGGGTGC  5242

Qy   2276  CCGCCCACAAGGGCATCGGCGGCAACGACAGCAGTGGCAAGGGCATCC  2335
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   5243  CCGCCCAACAAGGGCATCGGCGGCAACGACAGCAGTGGCAGGGGCATCC  5302

Qy   2336  GCAAGGTGCTGTTCCTGACGACGGCGGCATCGATGGGGGCATCGTGATCTACCAGTA  2387
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5303  GCAAGGTGCTGTTCCTGACGACGGCGGCATCGATGGGGGCATCGTGATCTACCAGTA  5354
```

```
RESULT 13
AX427926
LOCUS       AX427926                9194 bp    DNA        linear    PAT 20-JUN-2002
DEFINITION  Sequence 164 from Patent WO0232943.
ACCESSION   AX427926
VERSION     AX427926.1  GI:21538013
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 9194)
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 164 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9194
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN
Query Match     81.3%;  Score 2001.8;  DB 6;  Length 9194;
Best Local Similarity  91.3%;  Pred. No. 7e-201;
Matches  2166;  Conservative  0;  Mismatches  188;  Indels  19;  Gaps  4;
```

```
Qy   33    GGCCACCAGGCCAACATCTGATGCAGCGCAGCAACTTCAAGGGCCCAGCGCATCAT  92
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2989  GACCAACAGCGCCACCATCATGATGCAGCGCGGCAACTTCCGCAACCAGCCGCAAGATCGT  3048

Qy   93    CAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCGGCCCCCCGCAA  152
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3049  GAAGTGCTTCAACTGCGGCAAGGAGGGGCCACACCGCCCGCGGCCCCCCGCAA  3108

Qy   153   GAAGGGCTGCTGGAGTGCGGCAAGGAGGGCACCGAGGGGCGCCGCCA  212
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   3109  GAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCGAGGGGCGCCGCCA  3168

Qy   213   GGCCAACTTCTTCCCCGAGGACCTGGCCTTCCCCCAGGGCAGTTCCCCAG  272
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3169  GGCTAA-TTTTTAGGGAAGATCTGGCCTTCCCACAAGGGAGGGAATTTTCTTC  3227
```

```
Qy   902   TCCGGGAGCTGAACAAGCGGCACCCAGGACTTCTCTGGGAGGTGCAGCTGGCATCCCCACC  961
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3863  TCCGGGAGCTGAACAAGCGGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCACC  3922

Qy   962   CCGCCGGCTGAGAGAAGAGAGCGTGACCGGTGGCGACGTGCTGGACGCGCCTACTTCA  1021
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3923  CCGCCGGCCCTGAGAGAAGAGAGCGTGACCGGTGGCGACGTGCTGGACGCCTACTTCA  3982

Qy   1022  GGGTGCCCCTGACGACGTACACCCGCCTTCACCATCCCCAGCATCACA  1081
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3983  GCGTGCCCCTGACAAGGACTACACCCGCCTTCACCATCCCCAGCATCACA  4042

Qy   1082  ACGAGACCCCGGCATCGGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCC  1141
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4043  ACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCC  4102

Qy   1142  CCAGCATCTTCCAGAGCAGCAGCCAAGATCCTGGAGCCCTTCCGCGGCCCGCAACCCCG  1201
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4103  CCGCCATCTTCCAGTGCAGCATGACATCCGCCAAGATCCTGGAGCCCTTCCGCAAGCAGAACCCCG  4162

Qy   1202  AGATCGTGATCTACCA------GGCCCCCTGTACGTGGCAGCGACCTTGAGATCGGGC  1255
           |||||||||||||              |||||||||||||||||||||||||||||
Db   4163  ACATCGTGATCTACCAGTACATGGACGACCTGTACGTGGCAGCGACCTTGAGATCGGGC  4222

Qy   1256  AGCACCGGCCAAGATCGAGGAGCTGTGCGCAAGCACCTGCTCGTGCGCTGGGCTTCACCACCC  1315
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4223  AGCACCGGCCAAGATCGAGGAGCTGTGCGCAAGCACCTGCTCGTGCGCTGGGGGCTTCACCACCC  4282

Qy   1316  CCGACAAGAAGAGCACCAGAAGAAGAGCACCCCCTTCCTGTGGGATGGGCTACGAGCTGCACCCCCG  1375
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4283  CCGACAAGAAGAGCACCAGAAGAAGAGCACCCCCTTCCTGTGGGATGGGCTACGAGCTGCACCCCCG  4342

Qy   1376  ACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTTGGACCGTGAACGACA  1435
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4343  ACAAGTGGACCGTGCAGCCCATCGAGCTGCTGCTCGTCGTGCTGCCCGAGAAGGACCTGGACCGTGAACGACA  4402

Qy   1436  TCCAGAAGCTGGTGGGCAACTGACACCGGCCAGCCAGATCTACCCCGGCATCAAGGTGC  1495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4403  TCCAGAAGCTGGTGGGCAACTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGGTGC  4462

Qy   1496  GCCAGTGTGCAAGCTGCTGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCG  1555
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4463  GCCAGTGTGCAAGCTGCTGCGGCCGCACCCAAGGCCCTGACCGAGTGGTGCTGCCCCTGACCG  4522

Qy   1556  AGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGCCGTGCACGGGCG  1615
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4523  AGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGCCGTGCACGGGCG  4582

Qy   1616  TGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGAGACAGGCCACGACCAGT  1675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4583  TGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGAGACAGGCCAGT  4642

Qy   1676  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGA  1735
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4643  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGCA  4702

Qy   1736  TGCCGCACCGGCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCA  1795
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4703  TGAAGGGGCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCCCA  4762

Qy   1796  TGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCGCTGCCCATCCAGAAGGAGA  1855
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4763  CCGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAGGAGA  4822

Qy   1856  CCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTTGGGAGTTCG  1915
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4823  CCTGGGAGGCCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTTGGGAGTTCG  4882

Qy   1916  TGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG  1975
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4883  TGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG  4942
```

RESULT 14

```
AX427925
LOCUS       AX427925                9194 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 163 from Patent WO0232943.
ACCESSION   AX427925
VERSION     AX427925.1  GI:21538012
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 163 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9194
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match             81.2%;  Score 2000.2;  DB 6;  Length 9194;
Best Local Similarity   91.2%;  Pred. No. 1e-200;
Matches 2165;  Conservative 0;  Mismatches 189;  Indels 19;  Gaps 4;
```

```
Qy     33 GGCCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACTTCAAGGGCCCCAGCGCATCAT 92
          |||||| ||| |||||| ||| ||||||||||| ||  |||||||  |||||||||||||
Db   2989 GACCAACAGCGCCACCATCATGATGCAGCGCGGGCAACTTCCGCAACCAGCGCCAAGATCGT 3048

Qy     93 CAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCAACTGCGCCGCCCCCCGCCAA 152
          ||||||||||||||||||||||||||||||||  |  |||| |||||||||||||||||
Db   3049 GAAGTGCTTCAACTGCGGCAAGGAGGGCCAAGGAGGCCACCGCAACTGCGCCGCCCCCGCAA 3108

Qy    153 GAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAAGGAGGAGGCAAGGAGGACCCCCGCCA 212
          ||||||||||||||||||||||||||||||  |  |||||||||||||||||| |||||
Db   3109 GAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCAAGGAGGAGGCAAGGAGGGCCACGAGGACA 3168

Qy    213 GGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCAAGGCCCGGGAGTTCCCCAG 272
          ||||  ||||  |||| || |||||||| ||||||||||||||||||| |||||||||
Db   3169 GGCTAA-TTTTTAGGGAAGATCGGCCTTCCCCACAAGGGAAGGCCCAGGGAATTTTCTTC 3227

Qy    273 CGAGCAGAACCGCGCCACAGCGCCCAACAGCCGCCCACCAGCCGCGAGCTGCAGGTGCGCG------CGA 326
          |||||| ||||| ||||| |||||||||| ||||||  ||||||||||| || ||
Db   3228 AGAGCAGACCAGACCAACAGCCACCCACCCCCAGAGAGAGCTTCAGGTTTGGGGAGAGACAA 3287

Qy    327 CAACCCCCGCCAGCAGGCCGGCCGGCCGGCCGGCCAGGGCA------CCCTGAACTTCCCCCA 380
          ||||||| |||| |||| |||||  ||||  |||||||||
Db   3288 CAACTCCCTCTCAGAAGCAGGAGGCCGGATAGAAGAACTGTATCCTTTAGCTTCCTCTCA 3347

Qy    381 GATCACCCTGTGGCAGCGCGCCCCCTGGTGAGCATCAAGGTGGCGCGGCAGATCAAGGAGGC 440
          |||| |  |||||| || || || ||||| || || || |||| ||||||| ||||||
Db   3348 GATCACTCTTTGGCAGCGACACCCCTCGTCAGAAGGAATTAAAGATAGGGCCAGTCGAAGGAGGC 3407

Qy    441 CCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTG 500
          |||||||| || |||||| ||||| |||||| ||||| |||  ||||| |||||||||
Db   3408 CCTTCTAGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTG 3467

Qy    501 GAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGAT 560
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   3468 GAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGGGCCAGTACGACCAGAT 3527

Qy    561 CCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGT 620
          |||||||||||||||| ||||||| ||||||||||||||||||| ||||||||||||||
Db   3528 CCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGCTGGTGCCCCCCACCCCCGT 3587

Qy    621 GAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAG 680
          ||||||||||| ||||| ||||||||||  |||||||||||||||||||||||||||||
Db   3588 GAACATCATCGGCCGCGACATCGCTGCTGCCGGCGCTGCACCCTGCACCTTCCCCATCAG 3647

Qy    681 CCCCATCGAGACCGTGCCCGTGAAGCTGAAGAGCCGCATGGACGACGCCCAAGGTGAAGCA 740
          |||||||||||||||||||||||||||||||||||||
Db   3648 CCCCATCGAGACCGTGCCCGTGAAGCTGAAGAGCCGCATGGACGACGCCCAAGGTGAAGCA 3707

Qy    741 GTGGCCCTGACCGAGGAGAGAATCAAGGCCCTGACCGCCATCTGCGGAGGAGATGGAGAA 800
          |||||||||||||||||||| ||||||||||||  |||| ||||| ||||||||||||
Db   3708 GTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAAG 3767

Qy    801 GGAGGGCAAGATCACCAGAGATCGGCCCCGAGACACCCTACACACCCCGTTCGCCAT 860
          ||||||||||||| ||||||||||||||| || ||||| ||||||||||||| ||||
Db   3768 GGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCTACTACACCCCCGTTGTTCGCCAT 3827

Qy    861 CAAGAAGAGGACAGCACCAAGTGGCGCAAGCCAGTTCCGGACTTCCGCGAGCTGAACAAGCG 920
          |||| |||||||| ||||| |||| || ||  ||||| ||||||||||||| ||||||||
Db   3828 CAAGAGAGGACAGCACCAAGTGGCCAAGCTGGTCGAACTTCCGCGAGCTGAACAAGCG 3887

Qy    921 CACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCGGCCGGCCTGAAGAAGAA 980
          ||||||||||||  ||||||||||||| ||||||||||||| |||| || |||||||||
Db   3888 CACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCGGCCGGCCTGAAGAAGAA 3947

Qy    981 GAAGAGCGTGACAGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGAGGA 1040
          |||||| || |||| ||||||||||||||| |||||||||||||||||||||||| |||
Db   3948 GAAGAGCGGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGAGGA 4007

Qy   1041 CTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCG 1100
          ||||||||||||||||| ||||||||||||||||| |||||| |||||||||||||| ||
Db   4008 CTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCG 4067

Qy   1101 CTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCCAGAGCAG 1160
          |||||||||||||||||||||||||||| |||||| ||||||| ||||| ||||| |||
Db   4068 CTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCATCTTCTTCCAGTGCAG 4127

Qy   1161 CATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCCCGCCAAGATCGTGATCTACCA--- 1217
          ||||| ||||||||||||||||| |||||||| |||||||||||||||||| |||| |||
Db   4128 CATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACCCCGACATCGTGATCTACCAGTA 4187

Qy   1218 --GCCCCCCGTACGTGGGCGACCTGGAGCCAGCGACCGGCCAAGATCGA 1274
            ||| ||| || |||||||||||||||||| || ||| |||||||||
Db   4188 CATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGA 4247

Qy   1275 GGAGCTGGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGAGA 1334
          ||||||||| ||| || || || ||||||| || |||||||||| ||||| |||||||
Db   4248 GGAGCTGCGCCAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGAGA 4307

Qy   1335 GGAGCCCCCTTCGTGTGGAGAAGGAGCTGAAGGTGTGCGGCGG------CGA 1394
          ||||||||| |||  |  |||||| |||||||||||||||||||||||||
Db   4308 GGAGCCCCTTCCTGTGATGGGATGGTTCAGGTTTGGGGAGAGAGAGACAA 4367

Qy   1395 CATCGAGCTGCCCGAGAAGGAGAGCTGACCGTGAAGCAGTGGCCCTGCTGGAGCTGGTGGCAA 1454
          |||||| ||||| |||||||| |||  || |||||| || ||||| |||||||||||||
Db   4368 CATCGTGCTGTGCCCCAGAGGACAGCTGAACCGTGAACGAGCTGAACGAGCTGCCCAGCGTGGTGGCAA 4427

Qy   1455 GCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGATCAGCAAGAGCTGCT 1514
          || ||||| || || ||||| ||  ||||||||||||||||||||||||||||||||||
Db   4428 GCTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGGTGCGCCAGATCAAGCAGCTGCTGCT 4487

Qy   1515 GCGGGCGGCAAGCACCGGCCGCGAGATCTGCGCCGCGACATCGTGCCCGGAGCTGCGAGAGT 1574
          | ||||||| |||||||| ||||||| ||  |||| || |||||||||| ||||| ||||
Db   4488 GCGGGCGGCCACCACCGGCCGCGACATCGCCGCCGGCGTGTGCGCCGAGAGCTGCTGGAGT 4547

Qy   1575 GGCCGAGAACCGGCGAGATCCTCGCGACGCGGCTGTACTACACCCCAGCAA 1634
          ||||||||||||||||||||||||| ||||||| ||| || ||||| ||||
Db   4548 GGCCGAGAACCGGCGAGATCCTGAAGGACCGCGCACGCCGGTGTACTACGACCCCCAGCAA 4607

Qy   1635 GGACCTGGTGGCGGCAAGGAGGCCATCGGCACCGTGCTGATGGGCGACCTACCAGATCTTACCA 1694
          |||||| | |||||||||| || | ||||| || || || |||| |  | | ||||||||||||
Db   4608 GGACCTGATCGCGGCAGATCCAGAAGCAGGGCCAGTTGGACCTTGACCTACCAGATCTTACCA 4667

Qy   1695 GGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAGAGATGCGCACCGCCCACACCAA 1754
          |||| ||||  ||||||| |||||| | |||  || || || ||| ||||| |||||||||
Db   4668 GGAGCCCTTCAAGAACCTGAAGACCCAAGTAGCGCCCATGAAGGGCGGCCACACCAA 4727

Qy   1755 CGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGAGCATCGTGATCTG 1814
          ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   4728 CGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCACCGAGAGCATCGTGATCTG 4787
```

```
Qy    93  CAAGTGCTTCAACTGCCGCAAGGAGGGCCACATCGCCGCGACAGCTGCCGCAACTGCGGCCCCCGCAA    152
          |||||||||||||||| ||||||||||| ||||||  ||||||  ||||||||||||||||||||||||
Db  3049  GAAGTGCTTCAACTGCCGCAAGGAGGAGGGCCACACCGCCGCAACTGCGGCCCCCGCAA           3108

Qy   153  GAAGGGCTGCTGGAAGTGCGGCAAGGAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCA        212
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3109  GAAGGGCTGCTGGAAGTGCGGCAAGGAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCGACA    3168

Qy   213  GGCCAACTTCTTCCGGCAGAAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGA        272
          ||| ||    ||||  |||||||||||||||||||||||||||||||||||||||||
Db  3169  GGCTAA-TTTTTAGGGAAGATCGGCCTTCCCACAAGGAGGGCCAGGGAATTTTCTTC           3227

Qy   273  CGAGCAGAACCGGCCAACAGCCCACCAGCGCGGAGCTGCAGGTGCGCGG------CGA       326
          ||||||||||||||||||||||||||||||||| ||||| |||||| |
Db  3228  AGAGCAGACCACCAGCCAACAGCCCCACCAGCTTCAGGTTTGGGGAAGAGACAA              3287

Qy   327  CAACCCCCGCAGCGAGGCCGGCGAGGCCCAGGCA------CCCTGAACTTCCCCCA        380
          |||| |||||||||||||||||||||||| |||||       |||||||||||| ||
Db  3288  CAACTCCCTCTCAGAAGCAGGAGCCCATAGACAAGCAAGAACTGTATCCTTTAGTTCCCTCA     3347

Qy   381  GATCCACCCTGTGGCACCGCGCCCCTGGTGGACACCGTGCTGCAGGCGGCAGATCAAGGAGGC    440
          ||| ||  ||  ||||| |||||| |||||||| |||| ||||| ||||| ||||||||||||
Db  3348  GATCACTCTTTGGCAGCGACCCCTCGTCGACAACATAAAGATAGGGGGGCAACTGAAGGAGGC    3407

Qy   441  CCTGCTGGACACCGGCGCCCCGACGACACCGTGCTGGACGGCTTCATCAAGGTGGCCAGTAGGTG    500
          ||||||||||||||| |||||||||| |||| |||| ||||| |||||||||||| || ||||||
Db  3408  CCTTCTAGACACCGGCAAGGCCGACGACACCGTGCTGGAGGAGATGAACCTGCCCGGCAGGTGG     3467

Qy   501  GAAGCCCAAGATGATCGGCGGCAAGAGGCCATCAGCACCGTGCTGATCGGCCCCACCCCGT    560
          |||||||||||||||| |||||| |||||||||||||| ||||| ||||||||||||||||
Db  3468  GAAGCCCAAGATGATCGGCGGCGGCTTCATCAAGGTGGGCCAGTACCAGATCAGGGGCATCGG     3527

Qy   561  CCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCGT        620
          |||||||||||||| ||||| |||||||||||| ||||||||| ||||||||||||||||
Db  3528  CCTGATCGAGATCTGCGGCGCCACCGTGCTGATCGGCCCCACCCCGT                    3587

Qy   621  GAACATCATCGGCCGCAAACATGCTGACCCAGCTGGCTGCGACCCTGAACTTCCCCATCAG     680
          |||||||||||||||||| ||| |||||||||||| |||| ||| ||||||||||||||||
Db  3588  GAACATCATCGGCCGCAACCTGCTGACCCAGATCGGCTGCACCCTGAACTTCCCCATCAG       3647

Qy   681  CCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATCGAACGCCCGGCCGCCCAAGGTGAAGCA  740
          |||||||||||||||| ||||||||||||||||||| ||||  |||||||||||||||||||||||
Db  3648  CCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATCGAACGCCCGGCCGCCCAAGGTGAAGCA  3707

Qy   741  GTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCGCGAGGATGGAGAA          800
          |||||||||||||||||||||||||||||| |||| ||||| |||||||||||||||
Db  3708  GTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGGAGAA      3767

Qy   801  GGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGTTCGCCAT        860
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  3768  GGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGTTCGCCAT        3827

Qy   861  CAAGAAGAAGGACAGCACCAAGTGGCGCAAGTCGCCAGGCTTCCGCGAGCTGAACAAGCG     920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  3828  CAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGCGCAAGTCGCCGAGCTGAACAAGCG       3887

Qy   921  CACCCAGGACTTCTGCAGCTGGAGGTGCAGCTGGAGCATCCCCCCGCCGGCCTGAAGAAGAA    980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3888  CACCCAGGACTTCTGGAGGTGCAGCTGGAGCATCCCCACCCCGCCGGCCTGAAGAAGAA       3947

Qy   981  GAAGAGCGGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGGACGAGGAGA    1040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3948  GAAGAGCGGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGGACGAGGAGGA   4007

Qy  1041  CTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACGAGACCCCCGGCATCCG        1100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4008  CTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACGAGACCCCCGGCATCCG          4067

Qy  1101  CTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCCAGACAG      1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4068  CTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCATCTTCCAGTGCAG        4127
```

```
Qy  1815  GGGCAAGAGACCCCCAAGTTCCGCCTGCCCATCCAGGAGAAGGAGACCTGGGACGTGTGGAC    1874
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||| |||||
Db  4788  GGGCAAGAGACCCCCAAGTTCAAGCTGCCCATCCAGGAGAAGGAGACCTGGGAGGTGTGGAC     4847

Qy  1875  CGACTACTGCGAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCTGGT       1934
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4848  CGAGTACTGCGAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCTGGT         4907

Qy  1935  GAAGCTGTGGTACCAGTCTGGAGAAGGAGCCATCATCGGCCGCGAGACCTTCTACGTGGA      1994
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  4908  GAAGCTGTGGTACCAGTCTGGAGAAGGAGCCCATCATCGGCCGCGAGACCTTCTACGTGGA     4967

Qy  1995  CGGGCCGCCAACCGCGAGACCAAGATCGGCCAAGGCCCGGCTACGTGACCGCGGGGCCG       2054
          |||||||||||||||||||||||||||| || ||||||||||||||||||||| || |||
Db  4968  CGGGCCGCCAACCGCGAGACCAAGCTCGGGCAAGGCCCGGCTACGTGACCGGCGCCGCCG     5027

Qy  2055  GCAGAAGATCGTGAGCCTGACCGAGACCACCAACCACCGAGCTGCAGGCTGCCATCCA       2114
          ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  5028  CCAGAAGGTGGTGCCCCTGACCGAGACCACCAACCACCGAGAGCTGCAGGCGGCCATCCA      5087

Qy  2115  GCTCGGCCCTGCAGGACAGCAGCGCGAGGTGAACATGTGACCGAGCCAGTAGGCCCT      2174
          ||| ||||||||||||||||||||||||||||||| |||||||||||| ||||||||||
Db  5088  CCTGCCCCTGCAGGACAGCAGCGGCCTGGAGGTGAACATGTGAACCGACCAGCAGTGACCCT   5147

Qy  2175  GGGCATCATCCAGGCCCAAGAGCAGAAGAGGTGTACCTGAGCGGCATCTGGTGAACCAGATCATCGA  2234
          ||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db  5148  GGGCATCATCCAGGCCCGAGCAGCCCGACAGGTGTACCTGGCCTGAGCGGCATCGGCATCATCGA    5207

Qy  2235  GCAGCTGATCAAGAAGGAGAAAGTGTACCTGGGTGGTCTGGGCCCGCCACAAGGGCATCGG     2294
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  5208  GCAGCTGATCAAGAAGGAGAAAGTGTACCTGGCCTGAGCGGCATGGGGCATCGG            5267

Qy  2295  CGGCAACGAGCAGCAGATCGACAAGCTGGTGAGCGCAAGGCCATCCGCAAGGTGCTGTTCCTGGA   2354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5268  CGGCAACGAGCAGCAGATCGACAAGCTGGTGAGCGCCAAGGCCATCCGCAAGGTGCTGTTCCTGGA  5327

Qy  2355  CGGCATCGATGGGCCGGCCGCATCGTGATCTACCAGTA       2387
          |||||||| |
Db  5328  CGGCATCGACCAGCACCGGCCCAGGAGCACGAGAA              5360
```

RESULT 15
AX427927

```
LOCUS       AX427927           12411 bp    DNA        linear    PAT 20-JUN-2002
DEFINITION  Sequence 165 from Patent WO0232943.
ACCESSION   AX427927
VERSION     AX427927.1  GI:21538014
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 165 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..12411
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN
```

```
Query Match          81.2%;   Score 2000.2;  DB 6;  Length 12411;
Best Local Similarity   91.2%;   Pred. No. 9.6e-201;
Matches 2165;  Conservative   0;  Mismatches 189;  Indels 19;  Gaps  4;

Qy    33  GGCCACCAGCGCGCCAACATCCTGATGCAGCGCAGCAACTTCAAGGGCCCCAAGCGCGATCAT     92
          | ||||||||||||||||| ||| | |||||||||| ||||||| ||| ||| ||||| |||
Db  2989  GACCAACAGCGCGCCACCATCATGATGCAGCGCCAACTTCCGCAACCAGCGCCAAGATCGT      3048
```

```
Qy   1161  CATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCGAGATCGTGATCTACCA---  1217
           |||||||||||||||||||||||||||  ||||||||   |||| ||||||||||||
Db   4128  CATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGACACCCGACATCGTGATCTACCAGTA  4187

Qy   1218  ---GCCCCGTACGTGGGCAGCGACCTGGACATCGGCCAGCACCGCGCCAGAGATCGA  1274
              ||||| |||||||| ||||| ||||||| |||||||||||||||| ||||||||
Db   4188  CATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCCACCAGAGATCGA  4247

Qy   1275  GGAGCTGCGCAAGCACCTGCTCGCGTGGGCTTCACCACCCCGACAAGAAGCACCAGAA  1334
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   4248  GGAGCTGCGCGCCAGCACCTGCTCGCGTGGGCTTCACCACCCCGACAAGAAGCACCAGAA  4307

Qy   1335  GGAGCCCCTTCCTGTGGATGGGCTACAGAGCTGCACCCGACAAGTGACCGTGCAGCC  1394
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   4308  GGAGCCCCTTCCTGTGGATGGGCTACGGGCTGCACCCGACAAGTGACCGTGCAGCC  4367

Qy   1395  CATCGAGCTGCCCGAGAAGGAGAGCTGACCGTGAACGACATCCAGAAGCTGGTGGGCAA  1454
           ||||||||||||| ||||||||||||| ||||||||||||||||||||||||||||||
Db   4368  CATCGTGCTGCCCGAGAAGGAGACAGCTGGACATCCAGAAGCTGGTGGGCAA  4427

Qy   1455  GCTGAACTGGGCCAGCCAGATCTACCCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCT  1514
           ||||||||||| ||||||||||||||||||||||||||||| |||||||||||||| ||
Db   4428  GCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCT  4487

Qy   1515  GCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGGAGCT  1574
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   4488  GCGCGGCGCCAAGGCCCTGACCGAGTGGTGCCCCTGACCGAGGAGGCCGAGCTGGGAGCT  4547

Qy   1575  GGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAA  1634
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   4548  GGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAA  4607

Qy   1635  GGACTTGGTGGCCGAGATCCAGAGACAGGGCCCACGACTGGGACCTACCAGATCTACCA  1694
           ||||||||||||||||||||||||| |||||||||| |||||||||||||||||||||
Db   4608  GGACTTGATCGCCGAGATCCAGAGACAGGGCCCAGGGCCTACCAGATCTACCA  4667

Qy   1695  GGAGCCCTTCAAGAACCTGAAGACCGAAGATACGCAAGTACGCCCACCGCACCCACCAA  1754
           ||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||||
Db   4668  GGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCCATGAAGGGCGCCCACCACCAA  4727

Qy   1755  CGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAAGATCGTGATCTG  1814
           |||||||||||||||||||||||||| ||||||||||| |||||||||||||||||||
Db   4728  CGACGTGAAGCAGCTGACCGAGGCCGTGCTGACGAGAGATCGCCACCGAGAGACATCGTGATCTG  4787

Qy   1815  GGGCAAGACCCCAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGAC  1874
           ||||||||||||| |||||||| ||||||||||||||||||||| ||||||||||||
Db   4788  GGGCAAGACACCCCCAGTTCAAGTTCAAGCTGCCCATCCAGAGAGGAGACCTGGGAGACCTGGTGGAC  4847

Qy   1875  CGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGGT  1934
           ||||||| |||||||||||| |||||||||||||||||||||||||||||||||
Db   4848  CGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCTGT  4907

Qy   1935  GAAGCTGTGGTGTACCAGCTGCGAGGAAGGAGGAGGCCATCATCGGCGCCGAGACTTCTACGTGA  1994
           ||||||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   4908  GAAGCTGTGTGTACCAGCTGCGAGGAAGGAGAAGCCAGAAGACCAGCTGCAGGCCATCCA  4967

Qy   1995  CGGGCGCCGCCAACCGCGAGACCAAGATCGGCAGGCCGGCTACGTGACCGGGGGCCG  2054
           |||| |||| ||||| |||||||||||||||||||||||||||||||||||||||||
Db   4968  CGGGCGCCGCCAACCGCGAGACCAAGATCGGGCAAGGGCGGCTACGTGACCGGGCGGCG  5027

Qy   2055  GCAGAAGATCGTGAGCGTGACCGAGACCACCAGAGACCTGCAGGCCATCCA  2114
           |||||||| ||||||||||||||||| |||||||||||||||||||||||||
Db   5028  CCAGAAGGTGGTGCCCTGCCCTGACCGACCACCAACCAGAAGACCAGCTGCAGGCCATCCA  5087

Qy   2115  GCTGGCCCTGCAGGACAGCGGCAGCGGCAGCGAGGTGAACATCGTGACCGACCAGCCGCCCT  2174
           |||||||| ||||||||||||||||||| |||||||||| ||||||||||||||||||||||
Db   5088  CCTGGCCCTGCAGGACAGCGGCAGCGGCAGCGAGGTGAACATCGTGACCGACCAGCCGCCCT  5147

Qy   2175  GGGCATCATCCAGGCCCAGCCCGACAAGGCCGAGCTGGTGAACCAGATCATCGA  2234
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5148  GGGCATCATCCAGGCCCAGCCCGACAAGGCCGAGCTGGTGAGCTGCCAGATCATCGA  5207

Qy   2235  GCAGCTGATCAAGAAGGAGAAGGAGGAGGTGTACCTGGGCGTCGGGTGCCCGCCCCCACAAGGGCATCGG  2294
```

```
Db   5208  GCAGCTGATCAAGAAGGAGAAGGAGGTGTACCTGGGCGTCGGGTGCCCGCCCACAAGGGCATCGG  5267
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy          CGGCAACGAGCAGATCGACAAGTGGTGAGCAAGGGCATCCCGCGGCAAGGTGCTGTTCCTGGA  2354
           ||||||||| ||||||||||| || |||||||||||||||||||||||||||||||||||
Db   5268  CGGCAACGAGCAGGTGGACGGCCTGGTGGTGAGCGGCGGCAAGGTGCTGTTCCCTGGA  5327

Qy   2355  CGGCATCGATGGCGGCGCATCGTGATCTACCAGTA  2387
           |||||||||| ||||   ||||   |||||
Db   5328  CGGCATCGACAAGGCCCAGGAGAGCACGAGAA  5360
```

Search completed: April 10, 2004, 12:49:14
Job time : 6336.67 secs

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    April 10, 2004, 02:45:46 ; Search time 621 Seconds
                                       (without alignments)
                                       16849.133 Million cell updates/sec

Title:           US-09-610-313-31
Perfect score:   2463
Sequence:        1 gtcgagcgccaccatggccga........gggctagcaccggtgaattc 2463

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        3373863 seqs. 2124099041 residues

Total number of hits satisfying chosen parameters:       6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    N_Geneseq_29Jan04:*
              1: geneseqn1980s:*
              2: geneseqn1990s:*
              3: geneseqn2000s:*
              4: geneseqn2000as:*
              5: geneseqn2000bs:*
              6: geneseqn2002s:*
              7: geneseqn2000aas:*
              8: geneseqn2003bs:*
              9: geneseqn2003cs:*
             10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2463	100.0	2463	6	ABL39960	Abl39960 Synthetic
2	2455.4	99.7	2457	7	ACA03547	Aca03547 Synthetic
3	2455.4	99.7	2457	9	ADC13265	Adc13265 DNA of HI
4	2442.2	99.2	2469	6	ABL39959	Abl39959 Synthetic
5	2436.2	98.9	2457	6	ABL39961	Abl39961 Synthetic
6	2430.2	98.7	2457	7	ACA03548	Aca03548 Synthetic
7	2430.2	98.7	2457	9	ADC13266	Adc13266 DNA of HI
8	2422.6	98.4	2445	7	ACA03546	Aca03546 Synthetic
9	2422.6	98.4	2445	9	ADC13264	Adc13264 DNA of HI
10	2415.6	98.1	3930	9	ADC13230	Adc13230 DNA of HI
11	2414	98.0	3930	9	ADC13231	Adc13231 DNA of HI
12	2414	98.0	3930	9	ADC13232	Adc13232 DNA of HI
13	2414	98.0	5184	9	ACA03591	Aca03591 Synthetic
14	2414	98.0	5184	9	ADC13279	Adc13279 DNA of HI
15	2383.6	96.8	3531	9	ADC13234	Adc13234 DNA of HI
16	2382	96.7	3537	9	ADC13236	Adc13236 DNA of HI
17	2381	96.7	5145	7	ACA03521	Aca03521 Synthetic
18	2381	96.7	5145	9	ADC13233	Adc13233 DNA of HI
19	2371	96.3	3538	9	ADC13235	Adc13235 DNA of HI
20	2322.6	94.3	3607	7	ACA03551	Aca03551 Synthetic
21	2322.6	94.3	3607	9	ADC13269	Adc13269 DNA of HI
22	2322.6	94.3	3624	9	ACA03550	Aca03550 Synthetic
23	2322.6	94.3	3624	9	ADC13268	Adc13268 DNA of HI
24	2304.4	93.6	3597	7	ACA03549	Aca03549 Synthetic
25	2304.4	93.6	3597	9	ADC13267	Adc13267 DNA of HI
26	2142	87.0	2466	7	ACA03542	Aca03542 Synthetic
27	2142	87.0	2466	7	ACC78506	Acc78506 HIV p2Pol
28	2122.8	86.2	2472	7	ACA03543	Aca03543 Synthetic
29	2122.8	86.2	2472	7	ACC78507	Acc78507 HIV p2Pol
30	2115.2	85.9	2460	7	ACA03541	Aca03541 Synthetic
31	2115.2	85.8	2460	7	ACC78505	Acc78505 HIV p2Pol
32	2114.2	85.8	3564	7	ACC78488	Acc78488 HIV GagPo
33	2114.2	85.8	3564	7	ACC78489	Acc78489 HIV GagPo
34	2113.6	85.8	4716	9	ACA03522	Aca03522 Synthetic
35	2113.6	85.8	4716	9	ADC13238	Adc13238 DNA of HI
36	2109.8	85.7	3999	7	ACC78484	Acc78484 HIV GagCo
37	2108.2	85.6	3999	7	ACC78485	Acc78485 HIV GagCo
38	2108.2	85.6	3999	7	ACC78486	Acc78486 HIV GagCo
39	2108.2	85.6	5283	7	ACA03584	Aca03584 Synthetic
40	2108.2	85.6	5283	7	ACC78529	Acc78529 HIV TatRe
41	2108	85.6	4713	7	ACA03592	Aca03592 Synthetic
42	2108	85.6	4713	9	ADC13280	Adc13280 DNA of HI
43	2107.6	85.6	3462	9	ADC13237	Adc13237 DNA of HI
44	2082.2	84.5	3735	7	ACA03545	Aca03545 Synthetic
45	2082.2	84.5	3735	7	ACC78509	Acc78509 HIV p2Pol

ALIGNMENTS

```
RESULT 1
ABL39960
ID   ABL39960 standard; DNA; 2463 BP.
XX
AC   ABL39960;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Synthetic construct PR975YM SEQ ID NO:31.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO20204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Engelbrecht S,  Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 9; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
```

```
CC   therapy and DNA immunisation applications. ABJ39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 U; 0 Other;

Query Match        100.0%;  Score 2463; DB 6; Length 2463;
Best Local Similarity 100.0%;  Pred. No. 2.8e-297;
Matches 2463;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
Qy     1  GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCGCCAACATCCTGATGCAG    60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCGCCAACATCCTGATGCAG    60

Qy    61  CGCAGCAACTTCAAGGGCCCAAGGGGCCATCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CGCAGCAACTTCAAGGGCCCAAGGGGCCATCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGC   120

Qy   121  CACATGCCCGCCAACTGCGCGCCGCCCCCCCGCAAGAGGGCTGCTGGAAGTGCGGCAAGGAG   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CACATGCCCGCCAACTGCGCGCCGCCCCCCCGCAAGAGGGCTGCTGGAAGTGCGGCAAGGAG   180

Qy   181  GGCCACCAGATGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCCGCGAGGACCTGCGCC   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GGCCACCAGATGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCCGCGAGGACCTGCGCC   240

Qy   241  TTCCCCAGGGCAAGGCCCGGAGTTCCCCGACAACCCGCCCAGCAGCAGCAACAGCCCCACC   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TTCCCCAGGGCAAGGCCCGGAGTTCCCCGACAACCCGCCCAGCAGCAGCAACAGCCCCACC   300

Qy   301  AGCCGCGAGCTGCAGGTGCGGCGGCACAACCCCAGGCAGGGCCGGCCCGAGCGCCGGCCAG   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  AGCCGCGAGCTGCAGGTGCGGCGGCACAACCCCAGGCAGGGCCGGCCCGAGCGCCGGCCAG   360

Qy   361  GGCACCCTGAACTTCCCCCAGATCACCACCTGTGTGGACGCCCTGGTGAGCGGCTTCATCAAGGTG   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GGCACCCTGAACTTCCCCCAGATCACCACCTGTGTGGACGCCCTGGTGAGCGGCTTCATCAAGGTG   420

Qy   421  GGCGGCCAGATCAAGGAGGAGGGCCGGCGCCGCCGGCCGCCGAGGACACCGTGCTGTGGAGGAG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GGCGGCCAGATCAAGGAGGAGGGCCGGCGCCGCCGGCCGCCGAGGACACCGTGCTGTGGAGGAG   480

Qy   481  ATGAGCCTGCCCCGGCAAGTGGAAGCCAAGATGATCGGGCGCATCGGCGGCTTCATCAAG   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ATGAGCCTGCCCCGGCAAGTGGAAGCCAAGATGATCGGGCGCATCGGCGGCTTCATCAAG   540

Qy   541  GTGCGGCAGTACGACCAGATCCTGATCGAGAGATCTGCGGCAAGAGGGCCATCGGCACCGTG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GTGCGGCAGTACGACCAGATCCTGATCGAGAGATCTGCGGCAAGAGGGCCATCGGCACCGTG   600

Qy   601  CTGATCGGCCCCACCCCCGTGAACATCATCGGCGAAACATGCTGACCCAGCTGGGCTGC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTGATCGGCCCCACCCCCGTGAACATCATCGGCGAAACATGCTGACCCAGCTGGGCTGC   660

Qy   661  ACCCTGAACTTCCCCATCAGCCCCATCAGCCCCATCGAGAGTGCCCGTGAAGCTGTGGAAGCTGTGTCGGCACCGCATG   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  ACCCTGAACTTCCCCATCAGCCCCATCAGCCCCATCGAGAGTGCCCGTGAAGCTGTGGAAGCTGTGTCGGCACCGCATG   720

Qy   721  GACGGCCCAAGGTGAGCAGTGGCCCCACCCGTGGGCCCCGACGAGGAAGATCAAGGCCCTGACCGGC   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GACGGCCCAAGGTGAGCAGTGGCCCCACCCGTGGGCCCCGACGAGGAAGATCAAGGCCCTGACCGGC   780

Qy   781  ATCTCGGAGGAGATGGAGAGGCAAGAGGGCCAAGATCACCAAGAACCCCTAC   840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  ATCTCGGAGGAGATGGAGAGGCAAGAGGGCCAAGATCACCAAGAACCCCTAC   840

Qy   841  AACACCCCCGTGTTCGCCATCAAGAGAGAAGGACAGCACCACCGGCCCCGTGGCAAGCTGTGTGAC   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  AACACCCCCGTGTTCGCCATCAAGAGAGAAGGACAGCACCACCGGCCCCGTGGCAAGCTGTGTGAC   900

Qy   901  TTCCGCCGAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTCCGCCGAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC   960

Qy   961  CCCGCCGGCCGCTGCTGAGAGAAGAAGAGAGGCGTGACCGGTGCTGCTGGGCGACGTGGGGCGACGGCCTACTTC  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CCCGCCGGCCGCTGCTGAGAGAAGAAGAGAGGCGTGACCGGTGCTGCTGGGCGACGTGGGGCGACGGCCTACTTC  1020

Qy  1021  AGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCACGCCTTCACCATCCCAGCATCAAC  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  AGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCACGCCTTCACCATCCCAGCATCAAC  1080

Qy  1081  AACGAGACCCCGGCATCCGCTACCAGTACAGTACAAGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  AACGAGACCCCGGCATCCGCTACCAGTACAGTACAAGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140

Qy  1141  CCCAGCATCTTCCAGACGCAGCATGACCAAGATCCTGGAGCGCTTCCGCGCCCGCAACCCC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CCCAGCATCTTCCAGACGCAGCATGACCAAGATCCTGGAGCGCTTCCGCGCCCGCAACCCC  1200

Qy  1201  GAGATGTGATCTACCAGCCCCTGTACGTGGGCAGCGAGCCTGGAGTGCGGCAGCAGCAC  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GAGATGTGATCTACCAGCCCCTGTACGTGGGCAGCGAGCCTGGAGTGCGGCAGCAGCAC  1260

Qy  1261  CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCGTGGGGCTTCACCACCCCGAC  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCGTGGGGCTTCACCACCCCGAC  1320

Qy  1321  AAGAAGCACCAGAAGAGACCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  AAGAAGCACCAGAAGAGACCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG  1380

Qy  1381  TGGACCGTGCAGCCCATCGAGCTGCGCGGCGCAACTGGACGTGAACTGGGGCCGCATCAAGGTGCCGCCAG  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TGGACCGTGCAGCCCATCGAGCTGCGCGGCGCAACTGGACGTGAACTGGGGCCGCATCAAGGTGCCGCCAG  1440

Qy  1441  AAGCTGGTGGGCAAGTGAACTGGAACTGGTGGGCGGCGCCCGGCCCAGATCGTGCCCGATCAAGGTG  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  AAGCTGGTGGGCAAGTGAACTGGAACTGGTGGGCGGCGCCCGGCCCAGATCGTGCCCGATCAAGGTG  1500

Qy  1501  CTGTGCAAGCTGCTGCGCGGCGCCCAAGGCCCCGCTGCCCCTGACCGACCATCGTGCCCCGAGGAG  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CTGTGCAAGCTGCTGCGCGGCGCCCAAGGCCCCGCTGCCCCTGACCGACCATCGTGCCCCGAGGAG  1560

Qy  1561  GCCGAGCTGGAGCTGCGCGAGAACCGCGAGATCCTGCGCGGCCCGTGCACGGCGTGTAC  1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  GCCGAGCTGGAGCTGCGCGAGAACCGCGAGATCCTGCGCGGCCCGTGCACGGCGTGTAC  1620

Qy  1621  TACGACCCCAGCAAGGACGACTGTGGTGGCCGCCAGTGACCGAGATGCGC  1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  TACGACCCCAGCAAGGACGACTGTGGTGGCCGCCAGTGACCGAGATGCGC  1680

Qy  1681  TACCAGATCTACCAGGAGCCCTTCAAGAACTGAAGACCCGCCAAGATGCGGC  1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TACCAGATCTACCAGGAGCCCTTCAAGAACTGAAGACCCGCCAAGATGCGGC  1740

Qy  1741  ACCGCCACCACCAACGACTGCTGACGAGCAGCTGCGTGCAGAGATGCCCATGGAG  1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  ACCGCCACCACCAACGACTGCTGACGAGCAGCTGCGTGCAGAGATGCCCATGGAG  1800

Qy  1801  AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCGCGCTGCCCTGCCCATCCCAGAAGGAGACCTGG  1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCGCGCTGCCCTGCCCATCCCAGAAGGAGACCTGG  1860

Qy  1861  GAGACCTGGTGACCGGACTACTGGCAGGCCGATCCCGAGTTCGTGAAC  1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  GAGACCTGGTGACCGGACTACTGGCAGGCCGATCCCGAGTTCGTGAAC  1920

Qy  1921  ACCCCCCCGGTGAAGCTGTGTGTGTACCAGCTGGTACGGCCATCATCGGCGCCGAG  1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ACCCCCCCGGTGAAGCTGTGTGTGTACCAGCTGGTACGGCCATCATCGGCGCCGAG  1980

Qy  1981  ACTTCTACGTGACGTGGACGCGGCCGCGCCGGCCCAAGGCCGACTACGTG  2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  ACTTCTACGTGACGTGGACGCGGCCGCGCCGGCCCAAGGCCGACTACGTG  2040
```

The invention describes a synthetic polynucleotide encoding 2 or more immunogenic HIV polypeptides, where at least 2 of the polypeptides are derived from different HIV subtypes. The polynucleotide is useful for immunisation, generation of packaging cell lines, or production of HIV polypeptides. The polynucleotide and its encoded proteins are useful as immunogenic compositions or vaccines for generating humoral or cellular immune responses against HIV in a subject, or for inducing neutralising antibodies against HIV. The gene delivery vector comprising the polynucleotide is also useful for DNA immunisation of, or for generating an immune response (e.g. a humoral or cellular immune response) in, a subject such as a mammal, particularly a human. This sequence encodes a human immunodeficiency virus immunogenic peptide

```
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
SQ   Sequence 2457 BP; 564 A; 835 C; 758 G; 300 T; 0 U; 0 Other;

Query Match          99.7%;  Score 2455.4;  DB 7;  Length 2457;
Best Local Similarity 100.0%;  Pred. No. 2.4e-296;
Matches 2456;  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;

Qy      1 GTGACGCGCCACCATGGCCGAGGCCATGAGCGCCAGGCCACCATCCTGATGCAG       60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGACGCGCCACCATGGCCGAGGCCATGAGCGCCAGGCCACCATCCTGATGCAG       60

Qy     61 CGCAGCAACTTCAAGGGCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCAGCAACTTCAAGGGCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120

Qy    121 CACATCGCCCGCAACTGCCGCGCCCCCGCAGCGGCCAGGCCAACTGCGTGGAAGTGCGGCAAGAGAG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACATCGCCCGCAACTGCCGCGCCCCCGCAGCGGCCAGGCCAACTGCGTGGAAGTGCGGCAAGAGAG  180

Qy    181 GGCCACCAGATGAAGGACTGCACCGAGCGGCCAAGGCCCAGGCCAACTTCTTCCGGCGACCTGCGC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCCACCAGATGAAGGACTGCACCGAGCGGCCAAGGCCCAGGCCAACTTCTTCCGGCGACCTGCGC  240

Qy    241 TTCCCCAGGGCAAGGCCGAGTTCCCCACCGCCGAGAACGCGGCCAACAGCCCCACC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCCCCAGGGCAAGGCCGAGTTCCCCACCGCCGAGAACGCGGCCAACAGCCCCACC  300

Qy    301 AGCCGCGAGCTGCAGGTGCGGCGGCCACAACCCCCCAGCGGCGGCCGAGCGCCCAG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCCGCGAGCTGCAGGTGCGGCGGCCACAACCCCCCAGCGGCGGCCGAGCGCCCAG  360

Qy    361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCCCTGGTGAGCATCAAGGTG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCCCTGGTGAGCATCAAGGTG  420

Qy    421 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG  480

Qy    481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG  540

Qy    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG  600

Qy    601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC  660

Qy    661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG  720

Qy    721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC  780

Qy    781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC  840
```

```
Qy   2041 ACCGACCCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGAGACCGAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACCCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGAGACCGAC 2100

Qy   2101 CTGCAGGGCCATCAGCTGGCCCTGCAGGACAGCGGCAGCGGCGAGGTGAACATCGTGACCGACTGA 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGGCGAGGTGAACATCGTGACCGACAC 2160

Qy   2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGGAGCCGAGAGCGAGCTGGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGGAGCCGAGAGCGAGCTGGTG 2220

Qy   2221 AACCAGATCATCGAGCAGCTGATCAAGAGAGGAGGAGGTGTACCTGAGCTGGGTGCCCGCC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCTGATCAAGAGAGGAGGAGGTGTACCTGAGCTGGGTGCCCGCC 2280

Qy   2281 CACAAGGGCATCGGCGCGCAACGAGATCGACAACGAGGCGGCATCGCCGGCATCTGTGTGAGCAAGGGCCATCCGCAAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCGCGCAACGAGATCGACAACGAGGCGGCATCGCCGGCATCTGTGTGAGCAAGGGCCATCCGCAAG 2340

Qy   2341 GTGCTGTTCCTGGACGGCATCGATGCGGCATCGATTAAAAGCTTCCCCGGGGGCTAGCACATGGACGACCTG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCATCGATGCGGCATCGATTAAAAGCTTCCCCGGGGGCTAGCACATGGACGACCTG 2400

Qy   2401 TACGTGGGGCAGCCGCGGCGGGGCCCTAGGATCGATTCGATTAAAAGCTTCCCGGGGGCTAGCACCCGGTGAA 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGGGCAGCCGCGGCGGGGCCCTAGGATCGATTCGATTAAAAGCTTCCCGGGGGCTAGCACCCGGTGAA 2460

Qy   2461 TTC 2463
          |||
Db   2461 TTC 2463
```

RESULT 2
ACA03547

```
ID   ACA03547 standard; DNA; 2457 BP.
XX
AC   ACA03547;
XX
DT   22-MAY-2003 (first entry)
DE   Synthetic DNA ecoding immunogenic HIV peptide #30.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO200304657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 35; 262pp; English.
```

```
Qy   781                     |||||||||||||||||||||||||||||||||||
Db          ATCTGCGAGGAGATGGAGAGGAGGGCAAGATCACCACCAAGATCGGCCCCGAGAACCCCTAC 840

Qy   841  AACACCCCGTGTTCGCCATCAAGGAAGAAGGACAGCACCAAGTGGCCAAGCTGGTGGAC 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  AACACCCCGTGTTCGCCATCAAGGAAGAAGGACAGCACCAAGTGGCCAAGCTGGTGGAC 900

Qy   901  TTCCGCGAGCTGAACAAGGCGCACCCAGGACTTCTGGGAGGTGAGCGTGAGCTGGGCATCGGGGCCATCTGGGGCGCCTACGTG 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTCCGCGAGCTGAACAAGGCGCACCCAGGACTTCTGGGAGGTGAGCGTGAGCTGGCAGC 960

Qy   961  CCGCGCGGCCTGAAGAGAAGAGAGAAGCGGTGACCGTGCTGGACGTGGGCGACGCCTACTTC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CCCGCGCGGCCTGAAGAGAAGAGAGAAGCGGTGACCGTGCTGGACGTGGGCGACGCCTACTTC 1020

Qy  1021  AGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACGCGCCTTCACCATCCCAGCATCAAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  AGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACGCGCCTTCACCATCCCAGCATCAAC 1080

Qy  1081  AACGAGACCCCGGATCGCTACCAGTACACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  AACGAGACCCCGGATCGCTACCAGTACACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140

Qy  1141  CCCGCATCTTCCAGAGCAGCAGCAGATGACCAAGATCCTGGAGCCCTTCCGCGGCGCAACCCC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CCCAGCATCTTCCAGAGCAGCAGCAGATGACCAAGATCCTGGAGCCCTTCCGCGGCGCAACCCC 1200

Qy  1201  GAGATCGTGATCTACCAGGCCCCCTGTACGTGGCGGCAGCGAGATCGGCCAGCAC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GAGATCGTGATCTACCAGGCCCCCTGTACGTGGCGGCAGCGAGATCGGCCAGCAC 1260

Qy  1261  CGGCCAAGAATCGAGAGATCGAGGAGCTGCTTCGCGCAAGCACCACCCCCCGAC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  CGGCCAAGAATCGAGAGATCGAGGAGCTGCTTCGCGCAAGCACCACCCCCCGAC 1320

Qy  1321  AAGAAGCACCAGAAGAGAGAGCCCCCCTTCCTGTGCTGGTACGAGCTGCACCCCGACAAG 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  AAGAAGCACCAGAAGAGAGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1380

Qy  1381  TGGACCGTGCAGCTCGAGCTGCCCATCGAGCTGCCCGAGGAAGGAGGAGCTGAACGACATCCAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TGGACCGTGCAGCTCGAGCTGCCCATCGAGCTGCCCGAGGAAGGAGGAGCTGAACGACATCCAG 1440

Qy  1441  AAGCTCGTGGTGCAAGTGGCAAGCTGACTGAACTGACCCCGATCTACCCCGTGCGCGCAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  AAGCTCGTGGTGCAAGTGGCAAGCTGACTGAACTGACCCCGATCTACCCCGTGCGCGCAG 1500

Qy  1501  CTGTGCAAGCTGCTGCGCGGCGGCCAAGGCCCTGACCGACATCGTCGCCCTGACCGAGGAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CTGTGCAAGCTGCTGCGCGGCGGCCAAGGCCCTGACCGACATCGTCGCCCTGACCGAGGAG 1560

Qy  1561  GCCGAGTGGACCTGGCCGAGAACCGCGAGATCCTGCGCAAGGCCCTGCACGGCGTGTAC 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  GCCGAGTGGACCTGGCCGAGAACCGCGAGATCCTGCGCAAGGCCCTGCACGGCGTGTAC 1620

Qy  1621  TACGACCCCAGCAAGAACCTGGTGTGCGCGAGATCCAGAAGCAGGGCCACGACCAGTGGACC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  TACGACCCCAGCAAGAACCTGGTGTGCGCGAGATCCAGAAGCAGGGCCACGACCAGTGGACC 1680

Qy  1681  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1740

Qy  1741  ACCGCCCACAACAGCGACGTGAAGCAGCTGACAGAGGCCGTGCAGAAGATCGCCATGGAG 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  ACCGCCCACAACAGCGACGTGAAGCAGCTGACAGAGGCCGTGCAGAAGATCGCCATGGAG 1800

Qy  1801  AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAGGACCTGG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAGGACCTGG 1860

Qy  1861  GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1861  GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920

Qy  1921  ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAAGGAGACCCCATCATCGGGGCCGAG 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAAGGAGACCCCATCATCGGGGCCGAG 1980

Qy  1981  ACCTTCTACGTGGACGGGCGCGCCAACCGGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  ACCTTCTACGTGGACGGGCGCGCCAACCGGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040

Qy  2041  ACCGACCGGGGCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  ACCGACCGGGGCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2100

Qy  2101  CTGCAGGCCATCCAGTTGCCCTGCCCTGCCGCACGCGGCAGCAGGTGAACATCGTGACCGAC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  CTGCAGGCCATCCAGTTGCCCTGCCCTGCCGCACGCGGCAGCAGGTGAACATCGTGACCGAC 2160

Qy  2161  AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACACAGAGGAGGAGAGCTGGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACACAGAGGAGGAGAGCTGGTG 2220

Qy  2221  AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280

Qy  2281  CACAAGGGCATCGGCGGCAACGAACAGCAGATCGACCAAGCTGGTGAGCAAGGGCATCCCCAAG 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CACAAGGGCATCGGCGGCAACGAACAGCAGATCGACCAAGCTGGTGAGCAAGGGCATCCCCAAG 2340

Qy  2341  GTGCTGTTCCTGGACGGGCAGCATCGTGATCGTCGCCATCGTGATCTACCAGTACATGGACGACCTG 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GTGCTGTTCCTGGACGGGCAGCATCGTGATCGTCGCCATCGTGATCTACCAGTACATGGACGACCTG 2400

Qy  2401  TACGTGGGCAGCGGCGCGGCCCCTAGGGATCGATTAAAGCTTCCGGGGCTAGCACCGGT 2457
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  TACGTGGGCAGCGGCGCGGCCCCTAGGGATCGATTAAAGCTTCCGGGGCTAGCACCGGT 2457
```

RESULT 3
ADC13265
ID   ADC13265 standard; DNA; 2457 BP.
XX
AC   ADC13265;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct p2Pol-opt-YM_C SEQ ID NO 44.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,

PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 41; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2457 BP; 564 A; 835 C; 758 G; 300 T; 0 U; 0 Other;

Query Match   99.7%;   Score 2455.4; DB 9; Length 2457;
Best Local Similarity 100.0%;   Pred. No. 2.4e-296;
Matches 2456; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   1 GTCGAGGCCACCATGCCGAGGCCATGAGCCAGGCCACCAGCCCAACATCCTGATGCAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTCGAGGCCACCATGCCGAGGCCATGAGCCAGGCCACCAGCCCAACATCCTGATGCAG   60

Qy   61 CCCAGCAACTTCAAGGCGCCCAAGTGCTTCAACTGCGGCAAGGTGCGGCAAGGAGGGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGCAGCAACTTCAAGGGCGCCAAGTGCTTCAACTGCGGCAAGGTGCGGCAAGGAGGGGC 120

Qy   121 CACATCGCCGCAACTGCCGCCCCCCCGCAAGAAGGGCTGTGGGAAGTGCGGCAAGGAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACATCGCCGCAACTGCCGCCCCCCCGCAAGAAGGGCTGCTGGGAAGTGCGGCAAGGAG 180

Qy   181 GGCACCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCACCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCC 240

Qy   241 TTCCCCCAGGGCAAGGCCCGAGTTCCCCAGCAGAACCGCGCCAACAGCCCCACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCCCCAGGGCAAGGCCCGAGTTCCCCAGCAGAACCGCGCCAACAGCCCCACC 300

Qy   301 AGCCGCGGAGCTGCCAGGTGCGCGGCGCACAACCCGCCGAGGCCGGCCGGCCGGCCGCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCCGCGGAGCTGCCAGGTGCGCGGCGCACAACCCGCCGAGGCCGGCCGGCCGGCCGCAG 360

Qy   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG 420

Qy   421 GGCGGCCAGATCAAGGAGGCCCTGCTGACCGGCGCGACCACCGTGCTGGAGGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGCCAGATCAAGGAGGCCCTGCTGACCGGCGCGACCACCGTGCTGGAGGAG 480

Qy   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCAATCGGCGGCTTCATCAAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCAATCGGCGGCTTCATCAAG 540

Qy   541 GTGGCGCCAGTAGTACGACCAGATCCTGATCGAGCATGCGCGCAAGAAGGCCATCGCACCGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGCGCCAGTAGTACGACCAGATCCTGATCGAGCATGCGCGCAAGAAGGCCATCGCACCGTG 600

Qy   601 CTGATCGGCCGCCCCCACCCCCCGTGCCGGCCGCAACATGCTGAACCCAGCTGGGGCTGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCCGCCCCCACCCCCCGTGCCGGCCGCAACATGCTGAACCCAGCTGGGGCTGC 660

Qy   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720

Qy   721 GACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC 780

Qy   781 ATCTGCGAGGAGATGGAGAAGGAGGGGCCAAGATCACCAAGATCGGCCCCGAGAACCCTAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTGCGAGGAGATGGAGAAGGAGGGGCCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840

Qy   841 AACACCCCCGTGTTCGCCATCAAGAAGACCAGCACCAAGTGGCGCAAGGTGGTGGACAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCCGTGTTCGCCATCAAGAAGACCAGCAGCCAAGTGGCGCAAGGTGGTGGACAC 900

Qy   901 TTCCGGAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCCATCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCGGAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCCATCCCCAC 960

Qy   961 CCCGCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGACCGTGACCGTGGGGGCGCCTACTTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGACCGTGACCGTGGGGGCGCCTACTTC 1020

Qy   1021 AGGGTGCCCCTGGCAGACTTCCGCGAGGACTTCCGCCAAGTACACCGCCTTCACCATCCCCAGCATCAAC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGGGTGCCCCTGGACGAGGACTTCCGCGAGGACTTCCGCCAAGTACACCGCCTTCACCATCCCCAGCATCAAC 1080

Qy   1081 AACAGACCCCCGGCATCCCGTACCAGTACAACGTGCTGCCCCAGGGGCTGGAAGGGGCAGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACGAGACCCCCGGCATCCCGTACCAGTACAACGTGCTGCCCCAGGGGCTGAACGGGGCAGC 1140

Qy   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC 1200

Qy   1201 GAGATCGTGATCTACCAGGCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAAGCAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCAGGCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAAGCAC 1260

Qy   1261 CGGCGCCAAGATCGAGGAGTCGGCAAGCACCTGTGCGTGGGGCTTCACCACCCCCGAC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGCCAAGATCGAGGAGTCGGCAAGCACCTGTGCGTGGGGCTTCACCACCCCCGAC 1320

Qy   1321 AAGAAGCACCAGAAGGAGCCCCCTTCCTGTGAGTCCGCGAGAAGGAGAGCCGTGAACGACATCCAG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGAAGCACCAGAAGGAGCCCCCTTCCTGTGAGTCCGCGAGAAGGAGAGCCGTGAACGACATCCAG 1380

Qy   1381 TGGACCGTGCAGCCCATCGAGTCGCCGGCAACCCCGAGCGCGGAGGCCGGCCAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGTCGCCGGCAACCCCGAGCGCGGAGGCCGGCCAG 1440

Qy   1441 AAGCTGGTGCAGGCAAGCTGAACTGGGCCCAGATCTACCCCGGCATCAAGGTGCGGCCAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGCAGGCAAGCTGAACTGGGCCCAGATCTACCCCGGCATCAAGGTGCGGCCAG 1500

Qy   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy   1561 GCCGAGCTGGAGCTGGCGCGCAAGCTGCCCGAGAACCCGCGGAAGCTGCGGGAGCCGGCCGGCTGTAC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGCTGGAGCTGGCGCGCAAGCTGCCCGAGAACCCGCGGAAGCTGCGGGAGCCGGCCGGCGTGTAC 1620

Qy   1621 TACGACCCCAGCAAGGACCTGGTGGCCGAAGAGCAGGGCCAAGATCAGGAAGCAGCCAGTGGACC 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TACGACCCCAGCAAGGACCTGGTGGCCGAAGAGCAGGGCCAAGATCAGGAAGCAGCCAGTGGACC 1680

Qy   1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACTGGGGGCTGCAAGTACGCCAAGATCGCGC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACTGGGGGCTGCAAGTACGCCAAGATGTGCGC 1740

Qy   1741 ACCGCCCAACAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACCGCCCAACAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800

Qy   1801 AGCATCGTGATCTGGGGCAAGACCCCCCTGCCTCCATCCAGAGAGAGAACCTGG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCGTGATCTGGGGCAAGACCCCCCTGCCTCCATCCAGAGAGAGAACCTGG 1860

XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 8; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 U; 0 Other;

Query Match   99.2%; Score 2442.2; DB 6; Length 2469;
Best Local Similarity 99.6%; Pred. No. 1.1e-294;
Matches 2460;   Conservative 0;   Mismatches 3;   Indels 6;   Gaps 1;

```
Qy      1 GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCGAGCGCCAACATCCTGATGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCGAGCGCCAACATCCTGATGCAG 60

Qy     61 CGCAGCAACTTCAAGGGCCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGAGGGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCAGCAACTTCAAGGGCCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGAGGGGC 120

Qy    121 CACATGCCCGCAACTGCCGCCCCGCCCCCGCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACATGCCCGCAACTGCCGCCCCGCCCCCGCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAG 180

Qy    181 GGCCACCAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCGGCGAGGACCCGGCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCCACCAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCGGCGAGGACCTGGCC 240

Qy    241 TTCCCCAGGGCAAGGCCCGGGAGTTCCCCGACGAGCAGCAGCAACCGGCCAACAGCCCACC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCCCCAGGGCAAGGCCCGGGAGTTCCCCGACGAGCAGCAGCAACCGGCCAACAGCCCACC 300

Qy    301 AGCCGCGAGCTGCAGGTGCGGGACAACCCCGCAGCGAGGCCGGAGGCGGCCGGGGAGCCCAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCCGCGAGCTGCAGGTGCGGGACAACCCCGCAGCGAGGCCGGAGGCGGCCGGGGAGCCCAG 360

Qy    361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGGCCCCCCTGGTGAGCATCAAGGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGGCCCCCCTGGTGAGCATCAAGGTG 420

Qy    421 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCACCGGCGCCGGCCGACACCGAGCTGGAGGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCACCGGCGCCGGCCGACACCGAGCTGGAGGAG 480

Qy    481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540

Qy    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600

Qy    601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC 660

Qy    661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
```

```
Qy   1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCAGTGGAGGAGTTCGTGAAC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCAGTGGAGGAGTTCGTGAAC 1920

Qy   1921 ACCCCCCCTGGTGAAGCTGGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACCCCCCCTGGTGAAGCTGGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGGAG 1980

Qy   1981 ACCTTCTACGTGGACGGCGGCCGCCAACCCGGAGACCCGGAGACCGAGATCGGCAAGGCCGGCTACGTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACCTTCTACGTGGACGGCGGCCGCCAACCCGGAGACCCGGAGACCGAGATCGGCAAGGCCGGCTACGTG 2040

Qy   2041 ACCGACCGGGGCCGGCCAGAAGATCGTGAGCCTGCGGACCACCAACCAGAGACCGAG 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACCGGGGCCGGCCAGAAGATCGTGAGCCTGCGGACCACCAACCAGAGACCGAG 2100

Qy   2101 CTGCAGGCCATCCAGCTGGCCCTGAGGACAGCGGCAGCGAGTGAACATCGTGACCGAC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGTGAACATCGTGACCGAC 2160

Qy   2161 AGCCAGTACGCCCTGGCATCATCCAGGCCGACAAGAGCGAGAGCGAGCTGTGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCGACAAGAGCGAGAGCGAGCTGGTG 2220

Qy   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280

Qy   2281 CACAAGGGCATCGGCGGCCAACGAGCAGATCGACAAGCTGGTGAGCCAAGGGCATCCCGCAAG 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCGGCCAACGAGCAGATCGACAAGCTGGTGAGCCAAGGGCATCCCGCAAG 2340

Qy   2341 GTGCTGTTCCTGGACGGCGATCGATGGCGGCATCGTGATCTCACCAGTACATGACGACCTG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCGATCGATGGCGGCATCGTGATCTCACCAGTACATGACGACCTG 2400

Qy   2401 TACGTGGCAGCGCGGCGGGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGTTAGCACCGGT 2457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGCAGCGCGGCGGGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGTTAGCACCGGT 2457
```

RESULT 4
ABL39959
ID   ABL39959 standard; DNA; 2469 BP.
XX
AC   ABL39959;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic construct PR975(+) SEQ ID NO:30.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
FN   WO20020449493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.

RESULT 5
ABL39961
ID   ABL39961 standard; DNA; 2457 BP.
XX
AC   ABL39961;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic construct PR975YMWM SEQ ID NO:32.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO200204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX

PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 10; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABI39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 U; 0 Other;

Query Match   98.9%;   Score 2436.2;   DB 6;   Length 2457;
Best Local Similarity   99.6%;   Pred. No. 5.9e-294;
Matches 2454;   Conservative   0;   Mismatches   3;   Indels   6;   Gaps   1;

Qy   1 GTCGACGCGCCACCATGGCCAGGCCCATGAGCCAGGCCACCATCCTGATGCAG   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTCGACGCGCCACCATGGCCAGGCCCATGAGCCAGGCCACCATCCTGATGCAG   60

Qy   61 CGGAGCAACTTCAAGGGCCCCAAGGCCGATCATCAAGTGCTTCAACTGCGGCAAGAGGGC   120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGCAGCAACTTCAAGGGCCCCAAGGCCGATCATCAAGTGCTTCAACTGCGGCAAGAGGGC   120

Qy   121 CACATCGCCCGCAACTGCGCGCCCCGCGCCCCGCAAGAGAGGCTGCTGGAAGTGCCGCAAGGAG   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACATCGCCCGCAACTGCGCGCCCCGCGCCCCGCAAGAGAGGCTGCTGGAAGTGCGGCAAGGAG   180

Qy   181 GGCACCAGATGAAGGACTGCACCGAGGCCCAACTTCTTCGGCGAGGACCTGGCC   240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCCACCAGATGAAGGACTGCACCGAGGCGCCAACTTCTTCGGCGAGGACCTGGCC   240

Qy   241 TTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGAGAACCGGCCAAGCAGCCCCACC   300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGAGAACCGGCCAAGCAGCCCCACC   300

Qy   301 AGCCCGGAGCTGCAGGTGCGCGGCAACACCCCCGCAGGCGCCGAGGCCCGGCCAG   360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCCGCGAGCTGCAGGTGCGCGGCAACACCCCCGCAGGCGCCGAGGCCCGAGCGCCAG   360

Qy   361 GGCACCCTGAACTTCCCCAGATCACCCGTGCAGCGCCCCCTGTGAGCATCAAGGTG   420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCCACCCTGAACTTCCCCAGATCACCCGTGCAGCGCCCCCTGTGAGCATCAAGGTG   420

Qy   421 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGGCGAGACACCGTGTGGAGGAG   480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGCCAGATCAAGGAGGAGGCCCTGCTGGACACCGGCGACACCGTGTGCTGGAGGAG   480

Qy   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGGCTTCATCAAG   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCCGGCGCATGGAAGCCCAAGATGATCGGCGGCATCGGCGGGCTTCATCAAG   540

Qy   541 GTGCGCAGTACGACCACAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG   600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGCGCCAGTACGACCACAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG   600

Qy   601 CTGATCGGCGCCCACCCCCGTGAACATCATCGGCGCCGCAACATGCTGACCCAGCTGGGCTGC   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCGCCCACCCCCGTGAACATCATCGGCGCCGCAACATGCTGACCCAGCTGGGCTGC   660

Qy   661 ACCCTGAACTTCCCCATCAGCCCCATCAGCCCCATCGAGACGTGCCGGTGCGAGAGCTGAAGCTGAAGCCCGGCATG   720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCCATCAGCCCCATCAGCCCCATCGAGACGTGCCGGTGCGAGAGCTGAAGCTGAAGCCCGGCATG   720

Qy   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGAGAGATCAAGGCCTGACCGCC   780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGAGAGATCAAGGCCCTGACCGCC   780

Qy   781 ATCTGGAGGAGATGGAGAAGGAGGAGGGCAAGATCACCAAGATCACCCAGAACCCCTAC   840
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTGGAGGAGATGGAGAAGGAGGAGGGCAAGATCACCAAGATCACCCAGAACCCCTAC   840

Qy   841 AACACCCCGTGTTCGCCATCAAGAAGAAGCAGCACCAAGTGGGCCAAGCTGGTGGAC   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCGTGTTCGCCATCAAGAAGAAGCAGCACCAAGTGGGCCAAGCTGGTGGAC   900

Qy   901 TTCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCCAC   960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCGGCGTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCCAC   960

Qy   961 CCCCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTC   1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTC   1020

Qy   1021 AGGGTGCCCTGGACGAGGACTTCGGCAAGTACACCACCCGCCTTCACCATCCCAGCATCAAC   1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGGGTGCCCTGGACGAGGACTTCGGCAAGTACACCACCCGCCTTCACCATCCCAGCATCAAC   1080

Qy   1081 AACAGACCCCCGGCATCCGTACCAGTACAACGTGCTGCCCAGGGCTGAGAAGGGCCAGC   1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACAGAGACCCCCGGCATCCGTACACGGTACACAACGTGCTGCCCCAGGGCTGAAGGGCCAGC   1140

Qy   1141 CCCAGCATCTTCAGAGACAGCATGACCAAGATCCTGGAGCCCCTTCCGCGGCCCGCCAACCCC   1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCAGAGACAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCCGCCAACCCC   1200

Qy   1201 GAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCAGCGAGATCGGCCAGCAC   1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCAGCGAGATCGGCCAGCAC   1260

Qy   1261 CGCGGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGGCTTCACCACCCCCGAC   1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGGCTTCACCACCCCCGAC   1320

Qy   1321 AAGAAGCACCAGAAGGAGCCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG   1380
           |||||||||||||| |||||||||||||||
Db   1321 AAGAAGCACCAGAAGGAGCCCCCCCTTCCTGCCCAT-----CGGAGCTGCACCCCGACAAG   1374

Qy   1381 TGGACCGTGCAGCCCATCGAGCTGAACCTGGCCCGAGAGGAGGCTGACCGTGAACGACATCCAG   1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1375 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAGGAGGCTGACCGTGAACGACATCCAG   1434

Qy   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCCAG   1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1435 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCCAG   1494

Qy   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACGTGTGAGGAG   1560
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1495 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACGTGTGAGGAG   1554

Qy   1561 GCCGAGCTGGAGCTGGCCGAGAACCGGCAGATCCTGCGCGAGCCCGTGCACGGCGGCTTCATCAAG   1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1555 GCCGAGCTGGAGCTGGCCGAGAACCGGCAGATCCTGCGCGAGCCCGTGCACGGCGGCTTCATCAAG   1614

Qy   1621 TACCACCCCCAGGACGACCTGGTTGGCCAGGATCCAGAAGCAGAGGGCCACGACAGTTGGACC   1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1615 TACCACCCCCAGGACGACCTGGTTGGCCAGGATCCAGAAGCAGAGGGCCACGACAGTTGGACC   1674

Qy   1681 TACCAGATCTTACCAGGAGCGCCCTTCAAGAACCTGAAGACCGGCAAGTACGGCCAAGATCGCGC   1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1675 TACCAGATCTTACCAGGAGCGCCCTTCAAGAACCTGAAGACCGGCAAGTACGGCCAAGATCGCGC   1734

```
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 36; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 2457 BP; 568 A; 830 C; 758 G; 301 T; 0 U; 0 Other;

  Query Match           98.7%;  Score 2430.2;  DB 7;  Length 2457;
  Best Local Similarity 99.6%;  Pred. No. 3.3e-293;
  Matches 2448;  Conservative  0;  Mismatches  3;  Indels  6;  Gaps  1;

Qy      7 GCCACCATGGCCGAGGCCATGAGCCAGGCCAACATCCTGATGCAGGCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGAGGCCATGAGCCAGGCCAACATCCTGATGCAGGCAGC 60

Qy     67 AACTTCAAGGGCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGGGCAAGGGGCCACATC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGGGCAAGGGGCCACATC 120

Qy    127 GCCGCAACTGCCGCGCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCGCAACTGCCGCGCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC 180

Qy    187 CAGATGAAGGACTGCACCGAGCTGCACCGAGCGCCAGGCAGGCAGAACCGGCCCAACTTCCTCCGGCCTTCCCC 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAAGGACTGCACCGAGCTGCACCGAGCGCCAGGCAGGCAGAACCGGCCCAACTTCCTCCGGCCTTCCCC 240

Qy    247 CAGGGCAAGGCCCAGGCCCGGAGTTCCCCAGCGAGCACGAACCCGCGCCAACAGCCCCACCAGCGCACC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCAAGGCCCAGGCCCGGAGTTCCCCAGCGAGCACGAACCCGCGCCAACAGCCCCACCAGCGC 300

Qy    307 GAGCTGCAGGTGCGCGGCGACAACCCCCGAGCAGCGAGCCCCAGGAGCCGCCAGGGCACC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGCTGCAGGTGCGCGGCGACAACCCCCGAGCAGCGAGCCCCAGGAGCCGCCAGGGCACC 360

Qy    367 CTGAACTTCCCCAGATCACCCTGTGGCAGCATCAAGGTGGGCGGGC 426
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCAGATCACCCTGTGGCAGCATCAAGGTGGGCGGGC 420

Qy    427 CAGATCAAGGAGGAGGCCCTGCTGACGACACCGGCGCCGAGGAGATGAGC 486
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGGAGGCCCTGCTGACGACACCCGCGCCGAGGAGATGAGC 480
```

```
Qy   1741 ACCGCCCACCACCAACGACGTGAAGCAGCTGACCGACGGCCGTGCAGAAGATCGCCATGGAG 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1735 ACCGCCCACCACCAACGACGTGAAGCAGCTGAAGCAGCTGACGGCCGTGCAGAAGATCGCCATGGAG 1794

Qy   1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCGAGAGGAGACCTGG 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1795 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCGAGAGGAGACCTGG 1854

Qy   1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1855 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1914

Qy   1921 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1915 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1974

Qy   1981 ACTTCTACGTGACGGGCCCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1975 ACTTCTACGTGACGGGCCCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2034

Qy   2041 ACCGACCGCGGCCGGCAGAAGATCGTGACCGAGCCTGACCGAGACCAACCAGAAGACCGAG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2035 ACCGACCGCGGCCGGCAGAAGATCGTGACCGAGCCTGACCGAGACCAACCAGAAGACCGAG 2094

Qy   2101 CTGCAGGCCATCCAGCTGCCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2095 CTGCAGGCCATCCAGCTGCCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC 2154

Qy   2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCCGAGAGCGAGCTGGTG 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2155 AGCCAGTACGCCCTGGGCATCATCCAGGCCCCAGCCCGACAAGAGCCGAGAGCGAGCTGGTG 2214

Qy   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2215 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2274

Qy   2281 CACAAGGGCATCGGCGGCAACGAACAGCAGATCGACAAGCTGGTGAGCGCCAAGGGCATCCGCCAAG 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2275 CACAAGGGCATCGGCGGCAACGAACAGCAGATCGACAAGCTGGTGAGCGCCAAGGGCATCCGCCAAG 2334

Qy   2341 GTGCTGTTCCTGACGGCATCATGGCGGCATCTACCAGTACATGACGACGACCTG 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2335 GTGCTGTTCCTGACGGCATCATGGCGGCATCTACCAGTACATGACGACGACCTG 2394

Qy   2401 TACGTGGGCAGCGGCGGCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGTTCCCGGGGTGAA 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2395 TACGTGGGCAGCGGCGGCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGTTCCCGGGGTGAA 2454

Qy   2461 TTC 2463
          |||
Db   2455 TTC 2457
```

RESULT 6

ACA03548

```
ID   ACA03548 standard; DNA; 2457 BP.
XX
AC   ACA03548;
XX
DT   22-MAY-2003 (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #31.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO200300 4657-A1.
XX
PD   16-JAN-2003.
```

RESULT 7
ADC13266
ID   ADC13266 standard; DNA; 2457 BP.
XX
AC   ADC13266;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct p2Pol-opt_C SEQ ID NO 45.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX

OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,   Barnett SW,   Lian Y,   Engelbrecht S,   Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   Packaging cell lines.
XX
PS   Disclosure; Fig 42; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2457 BP; 568 A; 830 C; 758 G; 301 T; 0 U; 0 Other;

Query Match   98.7%; Score 2430.2; DB 9; Length 2457;
Best Local Similarity   99.6%; Pred. No. 3.3e-293;
Matches 2448; Conservative   0; Mismatches   3; Indels   6; Gaps   1;

Qy   7 GCCACCATGGCCGAGGCCATGAGCCAGGCCACCATCCTGATGCAGGCCAGC 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCCACCATGGCCGAGGCCATGAGCCAGGCCACCATCCTGATGCAGGCCAGC 60

Qy   67 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 126
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 120

Qy   127 GCCCGCAACTGCCGCCCCGCCCCGCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG 186
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCCGCAACTGCCGCGGCGCGCCCCGCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG 180

Qy   187 CAGATGAAGGACTGCACCGAGCGCGCCCCCGGTGCTGGAGCATCAAGTGGCGCGGC 246
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGATGAAGGACTGCACCGAGCGCGCCCCCGGTGCTGGAGCATCAAGTGGCGCGGC 240

Qy   247 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACGCAGAACCGCGCCAACCAGCGCC 306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACGCAGAACCGCGCCAACCAGCGCC 300

Qy   307 GAGCTGCAGGTGCGGCGGCGACAACCCCCGAGCGCGCGAGGGCCAGGGGCACC 366
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAGCTGCAGGTGCGGCGGCGACAACCCCCGAGCGCGCGAGGGCCAGGGGCACC 360

Qy   367 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCGCCCTGGTGAGCATCAAGTGGCGGGC 426
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCGCCCTGGTGAGCATCAAGTGGCGGGC 420

Qy   427 CAGATCAAGGAGGAGGCCCTGCTGGACACCGGCGCCGACACCGTGCTGGAGGAGATGAGC 486
           |||||||||||||||||||
Db   421 CAGATCAAGGAGGAGGCCCTGCTGGACACCGGCGCCGACACCGTGCTGGAGGAGATGAGC 480

Qy   487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC 540

Qy   547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGCCGCCACCGTGCTGATC 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGCCGCCACCGTGCTGATC 600

Qy   607 GGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG 666
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG 660

Qy   667 AACTTCCCCATCCAGCCCCATCGACACCCGTGCCCCGTGAAGCTGAAGCTGAAGTGACGGC 726
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AACTTCCCCATCCAGCCCCATCGACACCCGTGCCCCGTGAAGCTGAAGCTGAAGTGACGGC 720

Qy   727 CCCAAGGTGAAGCAGTGCCCCTGACCCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGC 786
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCCAAGGTGAAGCAGTGGCCCCTGACCCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGC 780

Qy   787 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 846
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 840

Qy   847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 906
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 900

Qy   907 GAGTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCACCCCGCC 966
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCCATCCCCACCCCGCC 960

Qy   967 GGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGCGACGCCTACTTCAGCGTG 1026
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGCGACGCCTACTTCAGCGTG 1020

Qy   1027 CCCCTGGACGAGGACTTCCGCAAGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCCAGCCCCAGC 1086
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGAGGACTTCCGCAAGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCCAGCCCCAGC 1080

Qy   1087 ACCCCCGGCATCCGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGGCAGCCCCAGC 1146
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCCCCGGCATCCGTCACGTCGCTGCCCCAGGGCTGGAAGGGGCAGCCCCAGC 1140

Qy   1147 ATCTTCCAGAGCCAGATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCTTCCAGAGCCAGATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1200

Qy   1207 GTGATCTACCA------GGCCCCCCGTGTACGTGGGCCAGGATCGGCAGCAC 1260
           |||||||||||
Db   1201 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCCAGGATCGGCAGCAC 1260

Qy   1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGAC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGAC 1320

Qy   1321 AAGAGGCACCCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGAGGCACCCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1380

Qy   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAG 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAG 1440

Qy   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAG 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAG 1500

Qy   1501 CTGTGCAAGCTGCTGCGGGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560
           ||||||

```
Db  1501 CTGTGCAAGCTGCTGCGCGGCGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560
Qy  1561 GCCGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGAGCCGTGCACGGCCGTGTAC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  1561 GCCGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGAGCCCGTGCACGGCCGTGTAC 1620
Qy  1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGACTCAGAAGCAGGGCCACGACCAGTGGACC 1680
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACC 1680
Qy  1681 TACCAGATCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATGCGC 1740
             ||||| |||||||||||||||  ||||||||||||||||||||||| || |||||||| |||
Db  1681 TACCAGATCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATGCGC 1740
Qy  1741 ACCGCCCCAACCACCACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800
             ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ACCGCCCACCACCACCACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800
Qy  1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860
Qy  1861 GAGACCTGGTGGACCGACTACTGCAGGCCACCTGGATCCCCGAGTTGGGAGTTCGTGAAC 1920
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 GAGACCTGGTGGACCGACTACTGCAGGCCACCTGGATCCCCGAGTTGGGAGTTCGTGAAC 1920
Qy  1921 ACCCCCCCGTGAAGCTGTGTACCAGCTGAGAAGGAGGCCCATCATCGGCGCCGAG 1980
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ACCCCCCCCGTGAAGCTGTGTACCAGCTGAGAAGGAGAGCCCATCATCGGCGCCGAG 1980
Qy  1981 ACCTTCTACGTGACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ACCTTCTACGTGACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
Qy  2041 ACCGACCGGGGCCGGCAGAGATCGTGAGCCTGACCGAGACCACCAGAAGACCGAG 2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ACCGACCGGGGCCGGCAGAGAGATCGTGAGCCTGACCGAGACCACCAGAAGACCGAG 2100
Qy  2101 CTGCAGGCCATCAGCTGCCCTGCAGGACAGCCGGGCAGGTGAACATCGTGACCGAC 2160
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGCAGGCCCATCCAGCTGCCCTGCAGGACAGCCGGGCAGCTGAACATCGTGACCGAC 2160
Qy  2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCCAGAGGTGGTG 2220
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTG 2220
Qy  2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db  2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280
Qy  2281 CACAAGGGCCATCGGCCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCCAAG 2340
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACAAGGGCCATCGGCCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2340
Qy  2341 GTGCTGTTCCTGGACGGCATCGATGCGGCATCGTGATCTACCAGTACATGGACGACCTG 2400
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GTGCTGTTCCTGGACGGCATCGATGCGGCATCGTGATCTACCAGTACATGGACGACCTG 2400
Qy  2401 TACGTGGGCAGCGACGGCGGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 TACGTGGGCAGCGACGGCGGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
```

RESULT 8
ACA03546
ID   ACA03546 standard; DNA; 2445 BP.
XX
AC   ACA03546;
XX
DT   22-MAY-2003 (first entry)
XX
DE   Synthetic DNA ecoding immunogenic HIV peptide #29.
XX

```
Query Match        98.4%;  Score 2422.6; DB 7;  Length 2445;
Best Local Similarity  99.6%;  Pred. No. 2.9e-292;
Matches 2441;  Conservative  0;  Mismatches  4;  Indels  6;  Gaps  1;

Qy    7 GCCACCATGGCCGAGGCCATGAGCCAGGCCAGGCCCAACATCCTGATGCAGCGCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GCCACCATGGCCGAGGCCATGAGCCAGGCCAGGCCCAACATCCTGATGCAGCGCAGC 60
Qy   67 AACTTCAAGGGCCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AACTTCAAGGGCCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 120
Qy  127 GCCCGCAACTGCCGCGCGCCCCGCCCCGCAAGAAGGGCCCAGGAAGTGGGCAAGGAGGGCCAC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GCCCGCAACTGCCGCGCGCCCCGCCCCGCAAGAAGGGCCCAGGAAGTGCGGCAAGGAGGGGCAC 180
Qy  187 CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 246
          ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db  181 CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 240
Qy  247 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACCCGCCCACCAGCCGC 306
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACCCGCCCACCAGCCGC 300
Qy  307 GAGCTGCAGGTGCGCGGCGACACCCCCGGCAGCCGCGGCCGCAGGGCACC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GAGCTGCAGGTGCGCGGCGACACCCCCGGCAGCCGCGGCCGCAGGGCACC 360
```

```
Qy   367 CTGAACTTCCCCCAGATCACCCTGGTGAGCATCAAGTGGCGGC 426
         ||||||||||||||||||||||||||||||||||||| ||||||
Db   361 CTGAACTTCCCCCAGATCACCCTGGTGAGCATCAAGTGGCGGC 420

Qy   427 CAGATCAAGGAGAGCCCTGCTGACACCGGCGCCACGACACCGGCGGAGATGAGC 486
         ||||||||||||||||||||||||||||||||||||||||||||
Db   421 CAGATCAAGGAGAGCCCTGCTGACACCGGCGCCACGACACCGGCGGAGATGAGC 480

Qy   487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGGGGCTTCATCAAGGTGCGC 546
         ||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGGGGCTTCATCAAGGTGCGC 540

Qy   547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC 606
         ||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC 600

Qy   607 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG 666
         ||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG 660

Qy   667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC 726
         ||||||||||||||||||||||||||||||||||||||||||||
Db   661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC 720

Qy   727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGC 786
         ||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGC 780

Qy   787 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 846
         ||||||||||||||||||||||||||||||||||||||||||||
Db   781 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 840

Qy   847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCCGGC 906
         ||||||||||||||||||||||||||||||||||||||||||||
Db   841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCCGGC 900

Qy   907 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCGCC 966
         ||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCGCC 960

Qy   967 GGCCTGAAGAAGAAGAGAGCGTGACCGTGAAGCGTGGGGGCGACCGCCTAGTTCAGCGTG 1026
         ||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCTGAAGAAGAAGAGAGCGTGACCGTGAAGCGTGGGGGCGACCGCCTAGTTCAGCGTG 1020

Qy  1027 CCCCTGGACGAGGACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCAACGAG 1086
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CCCCTGGACGAGGACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCAACGAG 1080

Qy  1087 ACCCCCGGCATCCGTACCAGTACAACGTGCTCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1146
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACCCCCGGCATCCGTACCAGTACAACGTGCTCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1140

Qy  1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCAACCCCGAGATC 1206
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCCAACCCCGAGATC 1200

Qy  1207 GTGATCTACCAGGCCCCCCTGTACGTGGCAGCGACCTGGCAGCACCGGCCAGTACACCGCGCC 1266
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTGATCTACCAGGCCCCCCTGTACGTGGCAGCGACCTGGCAGCACCGGCCAGTACACCGCGCC 1260

Qy  1267 AAGATCGAGGAGTGCGCAAGCACCCGCTTCACCACCCCCGACAAGAAG 1326
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AAGATCGAGGAGTGCGCAAGCACCCGCTTCACCACCCCCGACAAGAAG 1320

Qy  1327 CACCAGAAGGACCCCCCCTTCCTGTGGATGGACTACGAGCTGCACCCCCAGACAAGTGGACC 1386
         ||||||||||||||||||||||||
Db  1321 CACCAGAAGGAGCCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACAAGTGGACC 1374

Qy  1387 GTGCAGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTG 1446
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1375 GTGCAGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGACATCCAGAAGCTG 1434
```

```
Qy  1447 GTGGGCAAGCTGAACTGGGCCAGCCAGATCCCGGCATCAAGTGCGCCAGCTGTGC 1506
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1435 GTGGGCAAGCTGAACTGGGCCAGCCAGATCCCGGCATCAAGTGCGCCAGCTGTGC 1494

Qy  1507 AAGCTGCTGCGCGGCGGCGCCAAGGCCCTGACCGACACGTGCCCCTGACCGAGGAGGCCGAG 1566
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1495 AAGCTGCTGCGCGGCGGCGCCAAGGCCCTGACCGACACGTGCCCCTGACCGAGGAGGCCGAG 1554

Qy  1567 CTGGAGCTTGGCCAGAACCGCGAGATCCTGGGCGAGCCGTGCACGCGTGTACTACGAC 1626
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1555 CTGGAGCTTGGCCAGAACCGCGAGATCCTGGCGAGCCTGCACGGCGTGTACTACGAC 1614

Qy  1627 CCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGACCTACCAG 1686
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1615 CCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGACCTACCAG 1674

Qy  1687 ATCTACCAGGAGCCCTCAAGAACCTGAAGAACGGCAAGTACGCGCCAAGATGCGCACCGCC 1746
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1675 ATCTACCAGGAGCCCTTCAAGAACCTGAAGAACGGCAAGTACGCGCCAAGATGCGCACCGCC 1734

Qy  1747 CACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATC 1806
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1735 CACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATC 1794

Qy  1807 GTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAAGGAGACCTGGGAGAGACC 1866
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1795 GTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAAGGAGACCTGGGAGAGACC 1854

Qy  1867 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC 1926
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1855 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC 1914

Qy  1927 CCCCTGGTGAAGCTGTGTACAGCTGGAGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTC 1986
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1915 CCCCTGGTGAAGCTGTGTACAGCTGGAGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTC 1974

Qy  1987 TACGTGGACGGCGCCGCCAACCGCGAGACCAAGATGGCCAAGGCCCGCTACGTGACCGAC 2046
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1975 TACGTGGACGGCGCCGCCAACCGCGAGACCAAGATGGCCAAGGCCCGCTACGTGACCGAC 2034

Qy  2047 CGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG 2106
         ||||||||||||||||||||||||||||||||||||||||||||
Db  2035 CGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG 2094

Qy  2107 GCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAG 2166
         ||||||||||||||||||||||||||||||||||||||||||||
Db  2095 GCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAG 2154

Qy  2167 TACGCCCTGGACATCATCCAGGCGCGCCAGCCGGCCAGCTGGTGAACCAG 2226
         ||||||||||||||||||||||||||||||||||||||||||||
Db  2155 TACGCCCTGGACATCATCCAGGCGCGCCAGCCGGCCAGCTGGTGAACCAG 2214

Qy  2227 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAG 2286
         ||||||||||||||||||||||||||||||||||||||||||||
Db  2215 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAG 2274

Qy  2287 GGCATCGGCGGCAACGAGCAGATCGAGCAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTG 2346
         ||||||||||||||||||||||||||||||||||||||||||||
Db  2275 GGCATCGGCGGCAACGAGCAGATCGAGCAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTG 2334

Qy  2347 TTCCTGGACGGCGCGCATCGATGGCGGCATCCTGGATCCTCACCAGTACATGGACGACCTGTACGTG 2406
         ||||||||||||||||||||||||||||||||||||||||||||
Db  2335 TTCCTGGACGGCATCGATGGCGGCATCCTGGATCCTCACCAGTACATGGACGACCTGTACGTG 2394

Qy  2407 GGCAGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGTCTAGCACCGGT 2457
         ||||||||||||||||||||||||||||||||||||||||||||
Db  2395 GGCAGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGTCTAGCACCGGT 2445
```

```
AC   ADC13264;
XX
DT   18-DEC-2003  (first entry)
XX
DE   DNA of HIV construct p2Pol-opt-YMWM_C SEQ ID NO 43.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-03031192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Lian Y,  Engelbrecht S,  Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 40; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2445 BP; 562 A; 835 C; 751 G; 297 T; 0 U; 0 Other;

Query Match             98.4%;  Score 2422.6;  DB 9;  Length 2445;
Best Local Similarity   99.6%;  Pred. No. 2.9e-292;
Matches 2441;  Conservative    0;  Mismatches   4;  Indels   6;  Gaps   1;
```

```
Qy      7 GCCACCATGGCCGGAGGCCATGAGCGAGGCCAGGCCAACATCCTGATGCAGGGCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGGAGGCCATGAGCGAGGCCAGGCCAACATCCTGATGCAGGGCAGC 60

Qy     67 AACTTCAAGGGCCCCAAGCGCCCCCCCCGCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAAGGCTGCGGGAGTGCGGG 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCCAAGCGCCCCCCCCGCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAAGGCTGCGGGAGTGCGGG 120

Qy    127 GCCCCGCAACTGCCCGCCCCCCCCGCAAGGAGGCCGCGCAAGGAGGCTGCGGCCAAGGAGGGCCAC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCCGCAACTGCCCGCCCCCCCCGCAAGGAGGCCGCGCAAGGAGGCTGCGGCCAAGGAGGGCCAC 180

Qy    187 CAGATGAAGGACTGCACCGAGCGCCGGAGTTCCCCGGCAGCAGAACCGCGCCAAGACCGCGCCAGCCGC 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAAGGACTGCACCGAGCGCCGGAGTTCCTTCCCGCGGCAGCAGAACCGCGCCAAGACCGCGCCAGGTGCAC 240

Qy    247 CAGGGCAAGGCCCGGAGTTCCCCAGCGAGCAGAACCGCGCCAAGACCGCGCCACCAGCCGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCAAGGCCCGGAGTTCCCCAGCGAGCAGAACCGCGCCAAGACCGCGCCACCAGCCGC 300
```

```
Qy    307 GAGCTGCAGGTGCGCGGCGACAACCCCGCGAGCGAGGCCGCGCCGCCGAGGCCGCCCAGGGCACC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGCTGCAGGTGCGCGGCGACAACCCCGCGAGCGAGGCCGCCGCGCCGAGGCCGCCCAGGGCACC 360

Qy    367 CTGAACTTCCCCCAGATCACCCTGTGCAGCGCCCCGTGTGAGCATCAAGGTGGGCGGCGC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCCAGATCACCCTGTGCAGCGCCCCTGTGAGCATCAAGGTGGGCGGC 420

Qy    427 CAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGGAGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGC 480

Qy    487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC 540

Qy    547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC 600

Qy    607 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACCCTG 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACCCTG 660

Qy    667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGTGAAGCCCGGCATGGACGGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGTGAAGCCCGGCATGGACGGC 720

Qy    727 CCAAGGTGAAGCAGTGCCCCTGACCGGAGGAGAAGATCAAGGCCCTGACCGACCATCTGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCAAGGTGAAGCAGTGCCCCTGACCGGAGGAGAAGATCAAGGCCCTGACCGACCATCTGC 780

Qy    787 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 840

Qy    847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 900

Qy    907 GAGTGAACAAGGCACCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGTGAACAAGCGCACCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC 960

Qy    967 GGCCTGAAGAAGAAGGAGTCGGACGTGCTGGACGTGGGCGCACGCCTACTTCAGCGTG 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCTGAAGAAGAAGGAGTCGGACGTGCTGGACGTGGGCGCACGCCTACTTCAGCGTG 1020

Qy   1027 CCCCTGGACGAGCACTTCGGCAAGTACACCGCCTTCACCGCCCAGATCAACAACGAG 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGAGCACTTCGGCAAGTACACCGCCTTCACCGCCCAGATCAACAACGAG 1080

Qy   1087 ACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGCAGCCCCAGC 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGCAGCCCCAGC 1140

Qy   1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCGAGATC 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCGAGATC 1200

Qy   1207 GTGATCTACCAGGCCCCCCTGACCTGGGCAGCCAAGGAGATCGGCCAGCACCGCGCGCGC 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGATCTACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCGC 1260

Qy   1267 AAGATCGAGGAGCTGCGGCAAGCACCTGCTCTGGGGCTTCACCACCCCGAGAAGAAG 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AAGATCGAGGAGCTGCGGCAAGCACCTGCTCTGGGGCTTCACCACCCCGACAAGAAG 1320

Qy   1327 CACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGACAAGTGGACC 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CACCAGAAGGAGCCCCCCTTCCTGCCCAT-----CGAGCTGCACCCGACAAGTGGACC 1374

Qy   1387 GTGCAGCCCATCGAGCTGCCCGAGAAGGAGGAGCTGAACAGCTGACATCCGACACTG 1446
```

ADC13230
ID   ADC13230 standard; DNA; 3930 BP.
XX
AC   ADC13230;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagComplPolmut_C SEQ ID NO 9.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 6; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 890 A; 1365 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match   98.1%; Score 2415.6; DB 9; Length 3930;
Best Local Similarity   99.6%; Pred. No. 2e-291;
Matches 2434; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

```
Qy       14 TGGCGAGGCCATGAGCCAGCGACGGCCACCAGCGCCAACATCCTGATGCAGCGCGCAGCAACTTCA  73
            |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     1487 TCGCGGAGGCCATGAGCCAGCGACGGCCACCAGCGCCAACATCCTGATGCAGCGCGAGCAACTTCA 1546

Qy       74 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGGCGCAAGGAGGGCCACATCGCCCGCA  133
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1547 AGGGCCCCAAGCGCCATCATCATCAAGTGCTTCAACTGGCGCAAGGAGGGCCACATCGCCCGCA 1606

Qy      134 ACTGCCGCGCCCCCGCCCCCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCAAGAGGGGCAACCAGATGA  193
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1607 ACTGCCGCGCCCCCGCCCCCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCAAGAGGGGCAACCAGATGA 1666

Qy      194 AGGACTGCACGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA  253
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1667 AGGACTGCACGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 1726

Qy      254 AGGCCCCGCGAGTTCCCCAGCAGCAGAACCGGCGCAGAACCGGCCCACAGCCCCGCCGGAGCTGC  313
```

```
Db     1375                                                                         GTGCAGCGCCATCGAGCTGCCCGAGAAGAAGGAGAGCTGACCCGGCATCAAGGTGCGCCAGCTGTG 1434
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1447 GTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGAGAAGATCAAGGTGCGCCAGCTGTGC 1506
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1435 GTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGAGAAGATCAAGGTGCGCCAGCTGTGC 1494

Qy     1507 AAGTGCTGCTGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCGAG 1566
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1495 AAGTGCTGCTGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCGAG 1554

Qy     1567 CTGGAGCTGCGGCCGAGAACCGCGAGATCCTGCGCGAGCCGTGCACGGCCGTGTACTACGAC 1626
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1555 CTGGAGCTGCGGCCGAGAACCGCGAGATCCTGCGCGAGCCGTGCACGGCCGTGTACTACGAC 1614

Qy     1627 CCCAGCAACGACCTGGTGCCCAGGATCCAGAAGCAGGGCCACGACCAGTGACCTACCAG 1686
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1615 CCCAGCAAGGACCTGGTGCCCAGGATCCAGAAGCAGGGCCACGACCAGTGACCTACCAG 1674

Qy     1687 ATCTACCAGGAGACCCTTCAAGAACCTGAAGAACCGGCCAAGTACGCCCGAAGATGCGCACCGCC 1746
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1675 ATCTACCAGGAGACCCTTCAAGAACCTGAAGAACCGGCCAAGTACGCCCGAAGATGCGCACCGCC 1734

Qy     1747 CACACCAACGACGTGAAGCAGTTGACCGAGCCGTGCAGAAGATCGCCATGGAGAGCATC 1806
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1735 CACACCAACGACGTGAAGCAGCTGACCGAGCCCGTGCAGAAGATCGCCATGGAGAGCATC 1794

Qy     1807 GTGATCTGGGGCAAGACCCCCCAAGTTCGGCTTCGCCTGCCCATCCAGAAGGAGACCTGGGAGACC 1866
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db     1795 GTGATCTGGGGCAAGACCCCCAAGTTCGGCTTCGCCTGCCCATCCAGAAGGAGACCTGGGAGACC 1854

Qy     1867 TGGTGGACCCACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC 1926
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1855 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC 1914

Qy     1927 CCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGAAGGAGCCCATCATCGGCGCCGAGACCTTC 1986
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1915 CCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGAAGGAGCCCATCATCGGCGCCGAGACCTTC 1974

Qy     1987 TACGTGGACGGCGCCAACCGCGAGATCGGCAAGGCCGGCTACGTGACCGAC 2046
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1975 TACGTGGACGGCGGCGCCAACCGCGAGATCGGCAAGGCCGGCTACGTGACCGAC 2034

Qy     2047 CGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG 2106
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2035 CGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG 2094

Qy     2107 GCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAG 2166
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2095 GCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAG 2154

Qy     2167 TACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAG 2226
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2155 TACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAG 2214

Qy     2227 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAG 2286
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2215 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAG 2274

Qy     2287 GGCATCGGCGGCAACGAGGACGATGCGGCATCCGATGCTTACCAGGTGCTG 2346
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2275 GGCATCGGCGGCAACGAGGACGATGCGGCATCCGATGCTTACCAGGTGCTG 2334

Qy     2347 TTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTG 2406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2335 TTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTG 2394

Qy     2407 GGCAGCGGCGGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTTAGCACCGGT 2457
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2395 GGCAGCGGCGGCCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTTAGCACCGGT 2445
```

RESULT 10

```
Db  2807 AGGAGCCCCCTTCCTGCCCAT------CCGAGCTGCACCCGACCAAGTGACCGTGCAGC 2860
         ||||||||||||||||||||||      |||||||||||||||||||||||||||||||||
Qy  1394 CCATCGAGCTGCCCGAGAAGGAGAGCTGGAACGACATCGAGAGACTGTGGGCA 1453
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2861 CCATCGAGCTGCCCGAGAAGGAGAGCTGGAACGACATCGAGAGACTGTGGGCA 2920

Qy  1454 AGCTGAACTGGGCGCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC 1513
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2921 AGCTGAACTGGGCGCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC 2980

Qy  1514 TGCGCGGCGCCAAGGGCCCTGACCATCGTGCCCTGACCGAGGGCCGAGCTGGAGC 1573
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2981 TGCGCGGCGCCAAGGGCCCTGACCATCGTGCCCTGACCGAGGGCCGAGCTGGAGC 3040

Db  1907 TGCCGGAACCGCGAGATCTCGCGGCAGATCTCGCGGCAGATCCGTGTACTACGACCCCAGCA 1633
Qy  1574 TGGCCGGAACCGCGAGATCTCGCGGCAGATCTCGCGGCAGATCCGTGTACTACGACCCCAGCA 1633
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3041 TGGCCGAGAACCGCGAGATCCGCGAGATCTCGCGGCAGATCCGTGTACTACGACCCCAGCA 3100

Qy  1634 AGGACCTGGTGCCGAGATCCAGAAGCAGGGCCAGTGACCAGTGACCTACCAGATCTACC 1693
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3101 AGGACCTGGTGCCGAGATCCAGAAGCAGGGCCAGTGACCAGTGACCTACCAGATCTACC 3160

Qy  1694 AGGAGCCCTTCAAGAACCTGAGAACCTGAGAGATCGCGCCACACCA 1753
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3161 AGGAGCCCTTCAAGAACCTGAGAACCTGAGAGATGCGCCAAGATCGCGCCACACCA 3220

Qy  1754 ACGACGTGAAGCAGCTGACCGAGGCCGTGCCATCGTGAGAAGATCGTGATCT 1813
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3221 ACGACGTGAAGCAGCTGACCGAGGCCGTGCCATCGTGAGAAGATCGTGATCT 3280

Qy  1814 GGGCCAAGACCCCCCAAGTTCCGCTGCCCTGCCCATCCGAGAGGACCTGGGAGACCTGGTGGA 1873
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3281 GGGCCAAGACCCCCCAAGTTCCGCTGCCCTGCCCATCCGAGAGGACCTGGGAGACCTGGTGGA 3340

Qy  1874 CCGACTACTGCAGGCCACCTGGAGAAGAGCCCATCATCGGCGCGTACGTGACGGGGGCC 1933
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3341 CCGACTACTGCAGGCCACCTGGAGAAGAGCCCATCATCGGCGCGTACGTGACGGGGGCC 3400

Qy  1934 TGAAGCTGGTGTACCAGTGGAAGACCAAGATCGGCAAGGCCGGCTACGTGACGGGGGCC 1993
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3401 TGAAGCTGGTGTACCAGTGGAAGACCAAGATCGGCAAGGCCGGCTACGTGACGGGGGCC 3460

Qy  1994 ACGGGCGCGCCAACCGCGAGACCTGAGCCGAGAAGATCGGCCACCAACCAGAGACCAGGCC 2053
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3461 ACGGGCGCGCCAACCGCGAGACCTGAGCCGAGAAGATCGGCCACCAACCAGAGACCAGGCC 3520

Qy  2054 GGCAGAAGATCGTGACGGCCCAGGAGGTGAACATGGTGACCGAGCTGGTGAACCAGTACGCCC 2113
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3521 GGCAGAAGATCGTGACGGCCCAGGAGGTGAACATGGTGACCGAGCTGGTGAACCAGTACGCCC 3580

Qy  2114 AGCTGGCCCTGCACGAGCAGCCCAGCCCCGACACGCCCAGCCAGTAGCGCCC 2173
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3581 AGCTGGCCCTGCACGAGCAGCCCAGCCCCGACACGCCCAGCCAGTAGCGCCC 3640

Qy  2174 TGGGCATCATCCAGGCTGAAGGGCAGCCCAGCATCTTCC 2233
         |||||||||||||||||||||||||||||||||||||||
Db  3641 TGGGCATCATCCAGGCTGAAGGGCAGCCCAGCATCATCG 3700

Qy  2234 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGCCCCCACAAGGGGCATCG 2293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3701 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGCCCCCACAAGGGGCATCG 3760

Qy  2294 GCGGCAACGAGCAGCAGATCGGCCAGACCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGG 2353
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3761 GCGGCAACGAGCAGCAGATCGGCCAGACCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGG 3820

Qy  2354 ACGGGCATCGATGGCGCAAGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCCAGCG 2413
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3821 ACGGGCATCGATGGCGCAAGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCG 3880

Db  2747 GCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT 2457
Qy  2414 GCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
         |||||||||||||||||||||||||||||||||||||||||||||
Db  3881 GCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 3924
```

```
Db  1727 AGGCCCGGAGTTCCCAGCAGGAGCAGAACGCGCCAACAGCCCACCAGCCGCAGCTGC 1786
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   314 AGGCCCGGAGTTCCCAGCAGGAGCAGAACGCGCCAACAGCCCACCAGCCGCAGCTGAACT 373

Db  1787 AGGTGCGGCGGCGACAACCCCGCGAGCGAGGCGGCCGCGCCGAGCGCCAGGGCCACCCTGAACT 1846
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   374 AGGTGCGGCGGCGACAACCCCGCGAGCGAGGCGGCCGCGCCGAGCGCCAGGGCCAGATCA 433

Db  1847 TCCCCAGATCACCCTGTGCGAGCGCGCCCCCTGGTGACGCATCAAGGTGGGCGGCCAGATCA 1906
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   434 TCCCCAGATCACCCTGTGCGAGCGCGCCCCCTGGTGACGCATCAAGGTGGGCGGCCAGATCA 493

Db  1907 AGGAGGCCCTGCTGACACCGGCCGCCGACGACACCGGTGAGGATGAGCGTGGAGC 1966
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGCAACATGATCGGCGCAACATGATCGGCG 553

Db  1967 GCAAGTGGAAGCCCAAGATGATCGGCGCAACATGATCGGCGCAACATGATCGGCG 2026
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   554 ACCAGATCCTGATCGAGATCTGCGCGCAAGAAGGCCATGCGTGCTGCTGATCGGCGAGC 613

Db  2027 ACCAGATCCTGATCGAGATCTGCGCGCAAGAAGGCCATGCGTGCTGCTGATCGGCGAGC 2086
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   614 CCCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673

Db  2087 CCCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 2146
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   674 CCATCAGCGCCCCATCGCTGTGCCCGTGAAGCTGAAGCTGAAGCGGCCCCCAGG 733

Db  2147 CCATCAGCGCCCCATCGCTGTGCCCGTGAAGCTGAAGCTGAAGCGGCCCCCAGG 2206
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   734 TGGAGCAGTGGCCCCTGACCCAGGAGAAGATCAAGGCCCTGAACGCCGCCGAGGAA 793

Db  2207 TGGAGCAGTGGCCCCTGACCCAGGAGAAGATCAAGGCCCTGAACGCCGCCGAGGAA 2266
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   794 TCGCCATCAAGAAGAAGGACACCACCAAGATCGGCCCCCCGAGAACCCTACACACCCCGTGT 853

Db  2267 TCGCCCATCAAGAAGAAGGACACCACCAAGATCGGCCCCCCGAGAACCCTACACACCCCGTGT 2326
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   854 TCGCCATCAAGAAGAAGGACACCACCAAGATCGGCCCCCCGAGACTTCCGCGACTGA 913

Db  2327 TCGCCCATCAAGAAGAAGGACACCACCAAGATCGGCCCCCCGAGACTTCCGCGACTGA 2386
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   914 ACAAGGCCACCCGGACTTCTGGAGGTGCAGCTGGAGGTGCGAGTCCCCCGGCCTGA 973

Db  2387 ACAAGGCCACCCGGACTTCTGGAGGTGCAGCTGGAGGTGCGAGTCCCCCGGCCTGA 2446
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   974 AGAAGAAGAAGAGACTTCCGCAAGTACACCGGCCAAGGCACCACCCCCAGCATCAACAAGAGAGA 1033

Db  2447 AGAAGAAGAAGAGACTTCCGCAAGTACACCGGCCAAGGCACCACCCCCAGCATCAACAAGAGAGA 2506
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1034 ACGAGGACTTCCGCAAGTACACCGGCCAAGGCTTCACCACCCCCAGAAGCACCAGA 1093

Db  2507 ACGAGGACTTCCGCAAGTACACCGGCCAAGGCTTCACCACCCCCAGAAGCACCAGA 2566
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1094 GCATCCGCTACCAGTACACAAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 1153

Db  2567 GCATCCGCTACCAGTACACAAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 2626
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAATCGTGATCT 1213

Db  2627 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAATCGTGATCT 2686
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1214 ACCAGGCCCCCCGTACGGGGGCAGCGACCTGGGGCAGCGACCTGGAGATCGGGGGCACCCGGCACCGGCAAGATCG 1273

Db  2687 ACCAGGCCCCCCGTACGGGGGCAGCGACCTGGGGCAGCGACCTGGAGATCGGGGGCACCCGGCACCGGCAAGATCG 2746
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1274 AGGAGCTGCGCAAGCTGCCGCAAGTGGGCTGCTGCGGCTGCGGCTGCGGCTGCTGCGGCTGCGGCACCACCAGA 1333

Db  2747 AGGAGCTGCGCAAGCTGCCGCAAGTGGGCTGCTGCGGCTGCGGCTGCGGCTGCTGCGGCTGCGGCACCACCAGA 2806
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1334 AGGAGCTGCGCAAGCTGCCGCAAGTGGGCTGCTGCGGCTGCGGCTGCGGCTGCTGCGGCTGCGGCACCACCAGC 1393
```

RESULT 11
ADC13231
ID   ADC13231 standard; DNA; 3930 BP.
XX
AC   ADC13231;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagComplPolmutAtt_C SEQ ID NO 10.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide. The novel expression cassette can be used to treat HIV type C by gene therapy or
PT   packaging cell lines.
XX
PS   Disclosure; Fig 7; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 889 A; 1365 C; 1214 G; 462 T; 0 U; 0 Other;

Query Match   98.0%; Score 2414; DB 9; Length 3930;
Best Local Similarity 99.5%; Pred. No. 3.1e-291;
Matches 2433; Conservative 0; Mismatches 5; Indels 6; Gaps 1;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCGATGCAGCGCAGCAACTTCA 73
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCGATGCAGCGCAGCAACTTCA 1546

Qy   74 AGGGCCCCAAGCGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCA 133
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 AGGGCCCCAAGCGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCA 1606

Qy   134 ACTGCCGCGCCCCCGCAAGAAGGCTGCTGAAGTGCGGCAAGGAGGGGCCACCAGATGA 193
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 ACTGCCGCGCCCCCGCAAGAAGGCTGCTGAAGTGCGGCAAGGAGGGGCCACCAGATGA 1666

Qy   194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   1667 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 1726

Qy   254 AGGCCCGCGAGTTCCCCAGCAGGCAGAACCGGGCCAACAGCCCACCAGCCCGCGGGAGCTGC 313
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727 AGGCCCGCGAGTTCCCCAGCAGGCAGAACCGGGCCAACAGCCCACCAGCCCGCGGGAGCTGC 1786

Qy   314 AGTGCGGCGGCGACAACCCCGCAGGCGGCCGGCCGGCAGGCCACCCTGACT 373
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787 AGTGCGGCGGCGACAACCCCGCAGGCGGCCGGCCGGCAGGCCACCCTGACT 1846

Qy   374 TCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGTGGCGGCCAGATCA 433
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847 TCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGTGGCGGCCAGATCA 1906

Qy   434 AGGAGGCCCTGCTGACACACCCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG 493
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907 AGGAGGCCCTGCTGACACACCCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG 1966

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCGGCTTCATCAAGGTGCGCCAGTACG 2026

Qy   554 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGCCATGGCCACCGTGCTGATCGGCCCCA 613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 ACCAGATCCTGATCGAGATCTGCGGCAAGAGAGGCCATCGGCACCGTGCTGATCGGCCCCA 2086

Qy   614 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTTGGGCTGCACCCTGAACTTCC 673
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2087 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTTGGGCTGCACCCTGAACTTCC 2146

Qy   674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGGCGATGAAGCCGGCCCCAAGG 733
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2147 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGGCGATGAAGCCGGCCCCAAGG 2206

Qy   734 TGAAGCAGTGCCCCTGACCGAGAAGATCAAGGCCCTGACCGACGTGGCCATCTGCGAGGAGA 793
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2207 TGAAGCAGTGCCCCTGACCGAGAAGATCAAGGCCCTGACCGACGTGGCCATCTGCGAGGAGA 2266

Qy   794 TGGAGAAGGAGGGGCAAGATCACCAGGATCACCACCAGGATCGGCCCCGAGAACCCCGTGT 853
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2267 TGGAGAAGGAGGGGCAAGATCACCAGGATCACCACCAGGATCGGCCCCGAGAACCCCGTGT 2326

Qy   854 TCGCCCATCAAGAGAAGAGGACAGCACCAAGTGGCGCAAGTCGGTGGACTTCCGCGAGCTGA 913
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327 TCGCCCATCAAGAGAAGAGGACAGCACCAAGTGGCGCAAGTCGGTGGACTTCCGCGAGCTGA 2386

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGGCCTGA 973
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGGCCTGA 2446

Qy   974 AGAAGAAGAAGAGCGTGCTGGACGTGGGCGAGCTGCTCAGCGTGCCCCTG 1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2447 AGAAGAAGAAGAGCGTGCTGGACGTGGGCGAGCTGCTCAGCGTGCCCCTG 2506

Qy   1034 ACGAGGACTTCCGCAAGTACGGCTTCACCGCCTTCACCATCCCAGCATCAACAACGAGACCCCCG 1093
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2507 ACGAGGACTTCCGCAAGTACGGCTTCACCGCCTTCACCATCCCAGCATCAACAACGAGACCCCCG 2566

Qy   1094 GCATCCGCTACCAGTACAAGTGCTGCCCCAGGGCTGGAAGGGCCAGCCCCAGCATCTTCC 1153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2567 GCATCCGCTACCAGTACAAGTGCTGCCCCAGGGCTGGAAGGGCTGAAGGGCCAGCCCCAGCATCTTCC 2626

Qy   1154 AGAGCAGCATGACCAAGATCCAGAACTGCTGGGCAAGTGGGCGGCAAGGAGGGCCACATCGCCGCA 1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627 AGAGCAGCATGATGACCAAGATCCAGAACTGCTGGGCAAGTGGGCGGCAAGGAGGGCCACATCGCCGCA 2686

Qy   1214 ACCAGGCCCCCGTACGTGGGCAAGTGGCAGCGACCTGGAGATCGGCGCCAAGATCG 1273
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2687 ACCAGGCCCCCGTACGTGGGCAAGTGGCAGCGACCTGGAGATCGGCGCCAAGATCG 2746

Qy   1274 AGGAGCTGCGCAAGCACCTGCTGCGCTGGGCTTGCGCTGGGCTTGCGCTGGGCGAGCGAGAAGAAGCACCAGA 1333
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2747 AGGAGCTGCGCAAGCACCTGCTGCGCTGGGCTTGCGCTGGGCGAGCGAGAAGAAGCACCAGA 2806

```
Qy   2414 GCGGCCCTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT 2457
          |||||||||||||||||||||||||||||||||||||||||||
Db   3881 GCGGCCCTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT 3924
```

```
RESULT 12
ADC13232
ID   ADC13232 standard; DNA; 3930 BP.
XX
AC   ADC13232;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagCmp1Polmut1na_C SEQ ID NO 11.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200300460-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UVST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines
XX
PS   Disclosure; Fig 8; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 889 A; 1366 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match            98.0%;  Score 2414;  DB 9;  Length 3930;
Best Local Similarity  99.5%;  Pred. No. 3.1e-291;
Matches 2433;  Conservative 0;  Mismatches 5;  Indels 6;  Gaps 1;
```

```
Qy     14 TGGCCGAGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAACTTCA   73
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 TCGCCGGAGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAACTTCA 1546

Qy     74 AGGGCCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA  133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 AGGGCCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 1606

Qy    134 ACTGCCGCGCCCCCGCCCCCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA  193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 ACTGCCGCGCCCCCGCCCCCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA 1666
```

```
Qy   1334 AGGAGCCCCCTTCCTGTGATGGGCTACGAGCGTGCACCCCGACAAGTGGACCGTGCAGC 1393
          ||||||||||||||||                ||||||||||||||||||||||||||||
Db   2807 AGGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACAAGTGGACCGTGCAGC 2860

Qy   1394 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCA 1453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCA 2920

Qy   1454 AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCCAGCTGTGTGCAAGCTGC 1513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCCAGCTGTGTGCAAGCTGC 2980

Qy   1514 TGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCGAGCTGAGAGC 1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 TGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCGAGCTGAGAGC 3040

Qy   1574 TGGCCGAGACCGGCGAGATCCTGCGCGAGCCGTGCACGCGTGTACTACGACCCCAGCA 1633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 TGGCCGAGACCGGCGAGATCCTGCGCGAGCCGTGCACGCGTGTACTACGACCCCAGCA 3100

Qy   1634 AGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCCACGACCAGTCACCAGATCTACC 1693
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 AGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCCACGACCACCAGATCTACC 3160

Qy   1694 AGGAGCCCTTCAAGAACCTGAAGACGTACGGCAAGATGCGCACCGCCACCACA 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGCCCTTCAAGAACCTGAAGACCTGAAGATGCGCACCGCCACCACA 3220

Qy   1754 ACAGACCTGAAGCAGCAGTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCT 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 ACAGACCTGAAGCAGCAGTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCT 3280

Qy   1814 GGGCAAGAACACCCCCAAGTTCCGCCTCCCATCCGAGAAGGAGACCTGGTGGA 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 GGGCAAGAACACCCCCAAGTTCCGCCTCCCATCCGAGAAGGAGACCTGGTGGA 3340

Qy   1874 CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCCCTG 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCCCTG 3400

Qy   1934 TGAAGCTGTGGTACCAGTGGAAGGAGCCCATCATCGGCGCCGAGAACCTTCTACGTGG 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 TGAAGCTGTGGTACCAGTGGAAGGAGCCCATCATCGGCGCCGAGAACCTTCTACGTGG 3460

Qy   1994 ACGGCGCCAACCGCGAGATCCGAGACCAAGCGCCGGCTACGTGACCGACCGGGGCC 2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 ACGGCGCCAACCGCGAGATCCGAGACCAAGCGCCGGCTACGTGACCGACCGGGGCC 3520

Qy   2054 GGCAGAGATGTGAGCCTGACCGACCACCAACAGAGACCGAGCTGCAGGCCATCC 2113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 GGCAGAGATGTGAGCCTGACCGACCACCAACAGAGACCGAGCTGCAGGCCATCC 3580

Qy   2114 AGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCC 2173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 AGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCC 3640

Qy   2174 TGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGTGTGAACCAGATCATCG 2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 TGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGTGTGAACCAGATCATCG 3700

Qy   2234 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCGCCCACAAGGGATCG 2293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCGCCCACAAGAGGGATCG 3760

Qy   2294 GCGGCAACGAGCAGATCGACAAGCTGGTGAGCGAAGGGCATCGCCAAGGTGCGTTCCTGG 2353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 GCGGCAACGAGCAGATCGACAAGCTGGTGAGCGAAGGGCATCGCCAAGGTGCGTTCCTGG 3820

Qy   2354 ACGGGCATGCGATGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCG 2413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 ACGGCATCGATGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCG 3880
```

```
Qy   194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACTTGGCCTTCCCCCAGGGCA  253
         ||||||||||||||||||||||||| ||||||||||||||| ||||||||||||||||||
Db  1667 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 1726

Qy   254 AGGCCCGCGAGTTCCCAGCAGAACGCGCCAACAGCCCACCGCCGGCCGCGGCCGGAGCTGC  313
         ||||||||||||||||||||| ||||||||||||| ||||||||||||||||||||||||
Db  1727 AGGCCCGCGAGTTCCCAGCAGAACGCGCCGAGGCGGCCGGCCCCACCGCCGGCCGCGAGCTGC 1786

Qy   314 AGTGCGCGGCGACAACCCCCCGCAGCGGCGAGCGCCCGAGCGCCCGAGGCGCGGCACCCTGAACT  373
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1787 AGGTGCGCGGCGACAACCCCCCGCAGCGGCCGAGCGCCCGAGGCGCGGCACCCTGAACT 1846

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA  433
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1847 TCCCCCAGATCACCCTGTGGCAGCGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA 1906

Qy   434 AGAGGCCCTGCTGGACACGGCGGCATCGCCACCGCCGGCCGTGTACGAGGATAGAGGCCTGCCCG  493
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1907 AGAGGCCCTGCTGGACACGGCGGCATCGCCACCGCCGGCCGTGTACGAGGATAGAGGCCTGCCCG 1966

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGCGCCATCGACAAGGTGCGGCCAGTACG  553
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1967 GCAAGTGGAAGCCCAAGATGATCGGCGCGCCATCGACAAGGTGCGGCCAGTACG 2026

Qy   554 ACCAGATCCTGATCGACAGTCTGCGGGCCAACATGCTGACCCCAGTGCTGACTCGGCGACCCCCA  613
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2027 ACCAGATCCTGATCGACAGTCTGCGGGCCAACATGCTGACCCCAGTGCTGACTCGGCGACCCCCA 2086

Qy   614 CCCCGTGAACATCATCGGCCCAACATGCTGACCCCAGTGGGCTGCCATCTGACCCTGAAACTTCC  673
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2087 CCCCGTGAACATCATCGGCCCAACATGCTGACCCCAGTGGGCTGCCATCTGACCCTGAAACTTCC 2146

Qy   674 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCGCCATCTGCGAGGAGA  733
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2147 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCGCCATCTGCGAGGAGA 2206

Qy   734 TGGAGAAGGAGGGCCAAGATCACCAAGAATCACCCAGAAGATCGGCGCCCGAGAGATCGG  793
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2207 TGGAGAAGGAGGGCCAAGATCACCAAGAATCACCCAGAAGATCGGCGCCCGAGAGATCGG 2266

Qy   794 CTCGCCATCAAGAAGAAGCAGCACACCGTGGCACCACCGCCTTCGGCTGACTTCCGCGAGCTGTCG  853
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2267 TGGAGAAGAAGGGGCAAGATCACCAAGAATCACCCAGAAGATCGGCGCCCGAGAGATCGG 2326

Qy   854 TCGGCATCATCAAGAAGAAGACAGCACCACCGTGCCCCAGGGCCATCCCCCAGGGGCCTGAGCTGA  913
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2327 TCGGCATCATCAAGAAGAAGACAGCACCACCGTGCCCCAGGGCCATCCCCCAGGGGCCTGAGCTGA 2386

Qy   914 ACAAGGCGCACCCGAGGACTTCTGGGAGGTGCAGTTGCAGTTGCCCCACCCCCAGGGGCCTGA  973
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2387 ACAAGGCGCACCCGAGGACTTCTGGGAGGTGCAGTTGCAGTTGCCCCACCCCCAGGGGCCTGA 2446

Qy   974 AGAAGAAGAGAGCCTGCTGGACGTGGACGTGGACGTGGCGGACCCGCCTACTTCAGCGTCCCCCTGG 1033
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2447 AGAAGAAGAGAGCCTGCTGGACGTGGACGTGGACGTGGCGGACCCGCCTACTTCAGCGTGCCCCTGG 2506

Qy  1034 ACGAGGACTTCGCGAAGTACACCACCGGCAGTGCCGAGCATCCCAGCATCAACAACGAGACCCCCG 1093
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2507 ACGAGGACTTCGCGAAGTACACCACCGGCAGTGCCGAGCATCCCAGCATCAACAACGAGACCCCCG 2566

Qy  1094 GCATCCGCTACCAGTACAACACCGTGCTGCCCCAGGGGCTGGAAGGGGCAGCCCCAGCATCTTCC 1153
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2567 GCATCCGCTACCAGTACAACACCGTGCTGCCCCAGGGGCTGGAAGGGGCAGCCCCAGCATCTTCC 2626

Qy  1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTGCCCCGCGACCCCAAGCCCCGAGATCGTGATCT 1213
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2627 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCCGCGACCCCAAGCCCCGAGATCGTGATCT 2686

Qy  1214 ACCAGGCCCCCGTGTACGTGGGCGCAGCAGCGCGACCTGGACGCAGCCGAGATCGGCGCCGCCAAGATCG 1273
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2687 ACCAGGCCCCCGTGTACGTGGGCGCAGCAGCGCGACCTGGACGCAGCCGAGATCGGCGCCGCCAAGATCG 2746
```

```
Qy  1274 AGGAGCTGCCAAGCACCTGCTGCGCTGCGGCTTCACCCCCGACAAGAGCACCAGA 1333
         |||||||||||||||||||||||| |||||||||||||||||
Db  2747 AGGAGCTGCCAAGCACCTGCTGCGCTGCGGCTTCACCCCCGACAAGAGCACCAGA 2806

Qy  1334 AGGAGCCCCCTTCGTGTGGATGGGCTACGAGGCTGCAGTTGACCGTGCAGC 1393
         ||||||||||||||||| ||      |||||||||||||||||||
Db  2807 AGGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCGACCAAGTGACCGTGCAGC 2860

Qy  1394 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGACCGTGAACGACATCCAGAAGCTGTGGGCA 1453
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2861 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGACCGTGAACGACATCCAGAAGCTGTGGGCA 2920

Qy  1454 AGCTGAACTGGCCCAGCCAGCCCGATCTACCCCGGCATCAAGGTGCGCGCCAGCTGTGCAAGCTGC 1513
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2921 AGCTGAACTGGCCCAGCCAGCCCGATCTACCCCGGCATCAAGGTGCGCGCCAGCTGTGCAAGCTGC 2980

Qy  1514 TGCGCGGCGCCAAGGCCTGACCGACCACATCGTGCCCCGGCCGTGCGAGCCCGTGCGGAGGCCGAGC 1573
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2981 TGCGCGGCGCCAAGGCCTGACCGACCACATCGTGCCCCGGCCGTGCGAGCCCGTGCGGAGGCCGAGC 3040

Qy  1574 TGGCGAGACCGCGAGATCCGAGATCTCAAGGTGGTTCATCAAGGTGTGCTACTACGACCCCCAGCA 1633
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3041 TGGCGAGACCGCGAGATCCGAGATCTCAAGGTGGTTCATCAAGGTGTGCTACTACGACCCCCAGCA 3100

Qy  1634 AGGACCTGGTGGCCGAGATCCAGAGCAGGGCCACGACCAGCAGGCGCAAGAGTACGCCACCACCA 1693
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3101 AGGACCTGGTGGCCGAGATCCAGAGCAGGGCCACGACCAGCAGGCGCAAGAGTACGCCACCACCA 3160

Qy  1694 AGGAGCCTTCAAGAACCTGAAGACCGGCCGTGCAGAAGATCGGCCAAGTACGCCACCACCA 1753
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3161 AGGAGCCCTTCAAGAACCTGAAGACCGGCCGTGCAGAAGATCGGCCAAGTACGCCACCACCA 3220

Qy  1754 ACGACGTGAAGCAGCTGACCGTGAACCTGGGAGAGCATCGCGTGATCT 1813
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3221 ACGACGTGAAGCAGCTGACCGTGAACCTGGGAGAGCATCGCGTGATCT 3280

Qy  1814 GGGGCAAGACCCCAAGTTCCGCCTGCCATCCAGAAGAAGACCTGGGAGACCTGGTGA 1873
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3281 GGGGCAAGACCCCAAGTTCCGCCTGCCATCCAGAAGAAGACCTGGGAGACCTGGTGGA 3340

Qy  1874 CCGACTACTGGCAGGCCCACCTGGGATCCCCGAGTGGGAGTTCGTGAACACCCCCTGG 1933
         |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  3341 CCGACTACTGGCAGGCCCACCTGGGATCCCCGAGTGGGAGTTCGTGAACACCCCCTGG 3400

Qy  1934 TGAAGCTGTGTTACCAGCTGGACTTCGGCGGCCCCATCATCGGCGGCCCCTTCTACGTGG 1993
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3401 TGAAGCTGTGTTACCAGCTGGACTTCGGCGGCCCCATCATCGGCGGCCCCTTCTACGTGG 3460

Qy  1994 ACGGCCGCCGCCAACCGCGAGACCAAGATCGGCCAAGGCCCGGGCC 2053
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3461 ACGGCCGCCGCCAACCGCGAGACCAAGATCGGCCAAGGCCCGGGCC 3520

Qy  2054 GGCAGAAGAGATCGTGAGCGCTGACCGAGCTGCAGAGACCTGCGAGCTGCAGGCCATCC 2113
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3521 GGCAGAAGAGATCGTGAGCGCTGACCGAGCTGCAGAGACCTGCGAGCTGCAGGCCATCC 3580

Qy  2114 AGTGGCCCTGCAGGACAGCGGCGGAGGTGAACATCGTGACCAGCAGCCAGTAGCCCC 2173
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3581 AGTGGCCCTGCAGGACAGCGGCGGAGGTGAACATCGTGACCAGCAGCCAGTAGCCCC 3640

Qy  2174 TGGGCATCATCCAGGCCCGACAAGAGGCGAGCTGGTGAACCAGATCATCG 2233
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3641 TGGGCATCATCCAGGCCCGACAAGAGGCGAGCTGGTGAACCAGATCATCG 3700

Qy  2234 AGCAGTCGATCAAGAACCAGAGCAGATCGGTGTACCGAGCTGGTGCGCCCCAAGGGCATCG 2293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3701 AGCAGTCGATCAAGAACCAGAGCAGATCGGTGTACCGAGCTGGTGCGCCCCAAGGGCATCG 3760

Qy  2294 GCGGCAACCAGCAGCAGATCGACAAGAGCTGGTACGAGCAAGGGCATCGGCTGTTCCTGG 2353
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3761 GCGGCAACCAGCAGCAGATCGACAAGAGCTGGTACGAGCAAGGGCATCGGCTGTTCCTGG 3820

Qy  2354 ACGGCATCGATGGGCGCATCGTGATCGTACCAGTACCAGTGATCATGGACGACTCGTACGTGGCGGCAGCG 2413
```

```
Db   3821 ACGGCATCCATGGCGCGATCGTGATCTACCAGTACGATGGACGACCTGTACGTGGGCCAGCG 3880
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2414 GCGGCCCTAGGATCGATTAAAAGCTTCCGGGGCTAGCACCGGT 2457
          |||||||||||||||||||||||||||||||||||| ||||
Db   3881 GCGGCCCTAGGATCGATTAAAAGCTTCCGGGGCTAGCACCGGT 3924

RESULT 13
ACA03591
ID   ACA03591 standard; DNA; 5184 BP.
XX
AC   ACA03591;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Synthetic DNA ecoding immunogenic HIV peptide #74.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   immunisation, generation of packaging cell lines, or production of HIV
PT   polypeptides. The polynucleotide and its encoded proteins are useful as
PT   immunogenic compositions or vaccines for generating humoral or cellular
PT   immune responses against HIV in a subject, especially humans.
XX
PS   Example 1; Fig 79; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 5184 BP; 1139 A; 1852 C; 1610 G; 583 T; 0 U; 0 Other;

Query Match     98.0%;  Score 2414;  DB 7;  Length 5184;
Best Local Similarity   99.5%;  Pred. No. 3e-291;
Matches 2433;  Conservative    0;  Mismatches    5;  Indels    6;  Gaps    1;

Qy     14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCACCAGCATCCTGATGCAGCGCAACTTCA 73
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 TCGCCGAGGCCATGAGCCAGGCCACCAGCGCCACCAGCATCCTGATGCAGCGCAACTTCA 2800

Qy     74 AGGGCCCCAAGCCGATCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCCATCGCCCGCA 133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   2801 AGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCCATCGCCCGCCA 2860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   134 ACTGCGGCCGCCCCCCGAAGAAGGGCTCTGAGAGTGCGGCAAGAGAGGGCACCAGATGA 193
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   2861 ACTGCGGCCGCCCCCCGAAGAAGGGCTGCTGGAAGTGCGGCAAGAGAGGGCACCAGATGA 2920

Qy   194 AGGACTGCACCGAGCGGCCAACTTCTTCCGGGAGACCTGGCCTTCCCCCAGGGCA 253
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGACTGCACCGAGCGGCCAGCCAACTTCTTCCGGGAGACCTGGCCTTCCCCCAGGGCA 2980

Qy   254 AGGGCCGGAGTTCCCAGCGCAGAACGCGGCCAACAGCCCCAGCGCGGAGCTGC 313
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   2981 AGGCCCCGAGTTCCCAGCGCAGAACGCGGCCAACAGCCCCAGCGCGGAGCTGC 3040

Qy   314 AGGTGCGGCGACACCCCCGCAGGCGAGCGGCGGAGGGCCAGGCACCCTGAACT 373
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGGCGACAACCCCCGCAGGCGAGCGGCGGAGGGCCAGGCACCCTGAACT 3100

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGGTGAGCATCAAGTGGGGCGGCCAGATCA 433
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGGTGAGCATCAAGTGGGGCGGCCAGATCA 3160

Qy   434 AGGAGGGCCTGCTGACACCGGCGGCCGACGACACCGTGCTGGGAGGAGGATGAGCCTGCCCG 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGGCCCTGCTGCCGCCGCCGACGACACCGTGCTGGAGGAGGATGAGCCTGCCCG 3220

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAGGTTGCGGCCAGTACG 553
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   3221 GCAAGTGGAAGCCCAAGATCATCGGCGGCCTTCATCAAGGTTGCGGCCAGTACG 3280

Qy   554 ACCAGATCCTGATCGAGATCTCGGCAAGAGGCCATGGCCAGTGCTGCACCGTGCCCCA 613
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATCGAGATCTTGGCAAGAGGCCATGGCCAGTGCTGCACCGTGCCCCA 3340

Qy   614 CCCCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 3400

Qy   674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGGCCCAAGG 733
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGACGGGCCCAAGG 3460

Qy   734 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGTGCTGGACTTCGCGAGGAGA 3520

Qy   794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACACACCCCGTGT 853
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACACACCCCGTGT 3580

Qy   854 TCGCCATCAAGAAGAAGGACAGCAACAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCAAGAAGAAGGACAGCAACAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 3640

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCGCGGCCTGA 973
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCGCGGCCTGA 3700

Qy   974 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 3760

Qy   1034 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACAACGAGACCCCCG 1093
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACAACGAGACCCCCG 3820

Qy   1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAACGGCAGCCCCAGCATCTTCC 1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 3880

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCAAGGAGAGGGCCCATCGCCCGCA 1213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   3881 AGAGGAGCATGACCAAGATCCTGGAGCCCTTCCCGCGCCAACCCCGAGATCGTGATCT 3940

Qy   1214 ACCAGGCCCCCCTGTACGTGGGCAGCGCGACCTGGAGATCGGCGACCTGGAGATCG 1273
           ||||||| ||||||||||||||||||| ||||||||||||| |||||||||||||
Qy   2294 GCGGCAACGAGCGAGATCGACAAGCTGGTGAGCGAGGCATCGCAAGGCGTGCTTCCTGG 2353
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   5015 GCGGCAACGAGCGAGATCGACAAGCTGGTGAGCGAGGCATCGCAAGGCGTGCTTCCTGG 5074

Db   3941 ACCAGGCCCCCCTGTACGTGGGCAGCGCGACCTGGAGATCGGCGACCTGGAGATCG 4000

Qy   1274 AGGAGCTGCCCAAGCACCTGCTGCCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGA 1333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2354 ACGGCATCGATGCGGCATCGTGATCTTCACAGTACGTACATGGACGACCTGTACGTGGGCAGCG 2413
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5075 ACGGCATCGATGCGGCATCGTGATCTTCACAGTACGTACATGGACGACCTGTACGTGGGCAGCG 5134

Db   4001 AGGAGCTGGGCAAGCACCTGCTGCCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGA 4060

Qy   1334 AGGAGCCCCCTTCCTGTGATGGGCTACAGCTGCACCGTGACCGTGCAGC 1393
           |||||||||||| |||| |||||||||||| ||||| ||||||||||
Qy   2414 GCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
           |||||||||||||||||||||||||||||||||||||||||
Db   5135 GCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT 5178

Db   4061 AGGAGCCCCCCTTCCTTCCTGCCCAT------CGAGCTGCACCCGACAAGTGGACCGTGCAGC 4114

Qy   1394 CCATCGAGCTGCCCGAGAACGAGAGCATTGAACGACATCCAGAGAGCTGGTGGGCA 1453
           |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   4115 CCATCGAGCTGCCCGAGAACGAGAGCTGAACGACATCCAGAGAGCTGGTGGGCA 4174

Qy   1454 AGTCGAACTGGGCAGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCGACGCTGC 1513
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   14   TCGCCGAGGCCATGAGCCAGGCGCACCCAGCGCCAACATCCTGATGCCAGCAACTTCA 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4175 AGTCGAACTGGGCAGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCGACGCTGC 4234

Qy   1514 TGGCGCGGCGCCAAGGCCCTGACCGACGCCTGACCGGCCGAGCGGAGCTGGAGC 1573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4235 TGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGGCCGAGCTGGAGC 4294

Qy   1574 TGGCCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCA 1633
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4295 TGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCA 4354

Qy   1634 GGAGACCTGGTGCGCGAGATCCAGAAGCAGGGCCAGACCAGTGGACCTACCAGATCTACC 1693
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4355 AGGAGACCTGGTGCGCGAGATCCAGAAGCAGGGCCAGACCAGTGGACCTACCAGATCTACC 4414

Qy   1694 AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGAACGGCGCCCACACCA 1753
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4415 AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGAACGGCGCCCACACCA 4474

Qy   1754 ACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCT 1813
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4475 ACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCT 4534

Qy   1814 GGGGCCAAGACCCCCCAAGTTCCGCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGGA 1873
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4535 GGGGGCAAGACCCCCCAAGTTCCGCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGGA 4594

Qy   1874 CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGGAGTTCGTGAACACCCCCCCCTGG 1933
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4595 CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGGAGTTCGTGAACACCCCCCCCTGG 4654

Qy   1934 TGAAGCTGTGGTACCAGCTGAGAAGGAGCCCATCATCGGGCGCGCCGAGACCTTCTACGTGG 1993
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4655 TGAAGCTGTGGTACCAGCTGAGAAGGAGCCCATCATCGGGCGCGCCGAGACCTTCTACGTGG 4714

Qy   1994 ACCGGCGCGCCAACCGGCAGATCGGCGAGACCAAGATCGGCGAGACCTGGCAAGCGCCGGGGCC 2053
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4715 ACCGGCGCGCCAACCGGCAGATCGGCGAGACCAAGATCGGCGAGACCTGGCAAGCGCCGGGGGCC 4774

Qy   2054 GGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCC 2113
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4775 GGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCC 4834

Qy   2114 AGTTGGCCCTGCAGGACAGCGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACCCCC 2173
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4835 AGTTGGCCCTGCAGGACAGCGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCGCC 4894

Qy   2174 TGGGGCATCATCCAGGCCCGAGAGCGAGAGCCGAGAGCTGGTGAACCAGGATCATCG 2233
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4895 TGGGGCATCATCCAGGCCCAGCCCGAGAGCGAGAGCCGAGAGCTGGTGAACCAGGATCATCG 4954

Qy   2234 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCGGTGCCCGGCCCACAAGGGCATCG 2293
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4955 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGGCCCGCGGTGCCCGGCCCACAAGGGCATCG 5014

Db   2741 TCGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCGCAGCAACTTCA 2800
Qy   74 AGGGCCCCAAGCCGCATCATCATCAAGTGCTTCAACTGCGCGAAGAGGGCCACATCGCCCGCA 133
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AGGGCCCCAAGCCGCATCATCATCAAGTGCTTCAACTGCGCGAAGAGGGCCACATCGCCCGCA 2860

Qy   134 ACTGCCGGCGGCCCCCGCGCAAGAGAGGGCTGTGGAAGTGGGCAAGGAGAGGGCCACAGATGA 193
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 ACTGCCGGCGGCCCCCGCGCAAGAGAGGGCTGTGGAAGTGGGCAAGGAGAGGGCCACAGATGA 2920

Qy   194 AGGACTGCACCGACCGAGCGCGCCAGGCCAACTTCTTCGCCGAGGACCTGGCCTTCCCCCAGGGCA 253
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGACTGCACCGACCGAGCGCGCCAGGCCAACTTCTTCGCCGAGGACCTGGCCTTCCCCCAGGGCA 2980

Qy   254 AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCGCGCCCACCAGCCGCGAGCTGC 313
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCGCGCCCACCAGCCGCGAGCTGC 3040

Qy   314 AGGTGCGCGGCGACAACCCCCGCGAGGGCCAGGGCCAGGGCCACCGTGAACT 373
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGCGGCGACAACCCCCGCGAGGGCCAGGGCCAGGGCCACCGTGAACT 3100

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGCCGACGACACCGTGCTGGAGGAGATGAGGCTGCCCG 433
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCCAGATCACCCTGTGGCAGCGCCGACGACACCGTGCTGGAGGAGATGAGGCTGCCCG 3160

Qy   434 AGGAGGCCCTGCTGCACCGGCCCTGCTGGAGCAGACCGTGCTGCAAGGTGCGCCAGTACG 493
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGCCCTGCTGCACCGGCCCTGCTGGAGCAGACCGTGCTGCAAGGTGCGCCAGTACG 3220

Qy   494 GCAAGTGGAAGCCCAAGATGATGATCGGCGGCATCGGCGGCCATCGGCTGTGCTGCTGATCGGC 553
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATGATCGGCGGCATCGGCGGCCATCGGCTGTGCTGCTGATCGGC 3280

Qy   554 ACCAGATCTGATCGATGAGATCTGCGGCAAGAAGGCCATCGGCCACCGTGGATCGGCCCCA 613
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCTGATCGATGAGATCTGCGGCAAGAAGGCCATCGGCCACCGTGGATCGGCCCCA 3340

Qy   614 CCCCGTGAACATCATCGGCGCCGAACATGCTGGGCTGGGCTGCACCCTGAACTTCC 673
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCGTGAACATCATCGGCGCCGAACATGCTGGGCTGGGCTGCACCCTGAACTTCC 3400

Qy   674 CCATCAGCCCCATCGAGACCGACGACGTGCCCGTGAAGCTGAAGCCCGGCATGGACAGCGCCCAGG 733
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAGACCGACGACGTGCCCGTGAAGCTGAAGCCCGGCATGGACAGCGCCCAAGG 3460

Qy   734 TGAAGCAGTGCCCCGACCAGGAGATCAAGGCCCTGACCGCCATCTGCAGGAGA 793
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGGCCCCGACCAGGAGATCAAGGCCCTGACCGCCATCTGCAGGAGA 3520

Qy   794 TGGAGAAGAGGAGGGCAAGATCACCACCAGGGCCCCCGAGAACCCCTACAAACACCCCGTGT 853
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGAGAAGAGGAGGGCAAGATCACCACCAGGGCCCCCGAGAACCCCTACAAACACCCCGTGT 3580

Qy   854 TCGCCATCGAAGAAGAGGAGACAGCAGCACCAGCACCAGCGCCTACTTCCGGCGACTTCCCGGAGCTGA 913
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCGAAGAAGAGGAGACAGCAGCACCAGCACCAGCGCCTACTTCCGGCGACTTCCCGGAGCTGA 3640

Qy   914 ACAAGGCCACCAGGAGCTTCTGGGAGGTGCAGTCGGCGCATCGGCATCCCCACCCCGGCCTGA 973
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGGCCACCAGGAGCTTCTGGGAGGTGCAGTCGGCGCATCGGCATCCCCACCCCGGCCTGA 3700

Qy   974 AGAAGAGAGAGAGGAGGAGCTGACCGTGACCGTGCTGGACGTGGGGCGACCGCCTACTTCAGCGTGGGCGCCTGCTCAGCGTGGCCCTGG 1033
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAAGAGAGAGAGGAGGAGCTGACCGTGACCGTGCTGGACGTGGGGCGACCGCCTACTTCAGCGTGGGCGCCTGCTCAGCGTGGCCCTGG 3760

Qy   1034 ACGACGCCCTACCAGTACACAACGCGTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACACGACCCCCG 1093
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACGACGCCCTACCAGTACACAACGCGTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACACGAACCCCCG 3820

Qy   1094 GCATCCGCTACCGGTGTGCTGCCCGGGCTGGAAGGGCAGCCCCAGCATCTTCC 1153
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGCTACCGGTGTGCCCCGGGCTGGAAGGGCAGCCCCAGCATCTTCC 3880

Db   1154 AGAGCAGCATGACCAAGATCCTGAGCGCGCCGACCCTTCCCGCGCCCGCACCCCGAGATCGTGATCT 1213
Qy   3881 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCCGCGCCCGCACCCCGAGATCGTGATCT 3940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1214 ACCAGGCCCCCCGTGTACGTGGCAGCGACCTGGCAGGATCGGCAGTGAGATCGGCCCAAGATCG 1273
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3941 ACCAGGCCCCCCGTGTACGTGGCAGCGACCTGGCAGGATCGGCAGTGAGATCGGCCCAAGATCG 4000

Qy   1274 AGGAGCTGCGCAAGCACCTGCTGCTGGGCTTCACCACCCCCGACAAGCAGAAGACCACCAGA 1333
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4001 AGGAGCTGCGCAAGCACCTGCTGCTGGGCTTCACCACCCCCGACAAGCAGAAGACCACCAGA 4060

Qy   1334 AGGAGCCCCCTTCCTGTGGATGGCTACGAGCTGCACCCGACCGTGCAGC------- 1393
         ||||||||||||||||||||||||||||||||||| ||||||||||
Db   4061 AGGAGCCCCCTTCCTGCTGCCCAT------CGAGCTGCACCCGACCGTGCAGC 4114

Qy   1394 CCATCGAGCTGCCCGAGAAGAGAGTGGACCTGGAACGACATCCAGAAGCTGTGGGCA 1453
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4115 CCATCGAGCTGCCCGAGAAGAGAGTGGACCTGGAACGACATCCAGAAGCTGTGGGGCA 4174

Qy   1454 AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC 1513
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4175 AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC 4234

Qy   1514 TGCGGCGCGCCAAGGCCCTGACGCACATCGTGCCCCCTGCCACGGGCCGAGCTGGAGC 1573
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4235 TGCGGCGCGCCAAGGCCCTGACGCACATCGTGCCCCCTGCCACGGGCCGAGCTGGAGC 4294

Qy   1574 TGGCCGAGAACCGGCAGATCCTGCGCCAGGTGTACTACGACCCCAGCA 1633
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   4295 TGGCCGAGAACCGGCAGATCCTGCGCCAGGTGTACTACGACCCCAGCA 4354

Qy   1634 AGGACCTGGTGGCCAAGAAGCAGGAGGGCCAAGTCGCCCACCAGTGGACCAGATCTACC 1693
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4355 ACCAGATCTGGTGGCCAAGAAGCAGGAGGGCCAAGTCGCCCACCAGTGGACCAGATCTACC 4414

Qy   1694 AGGAGCCCTTCAAGAACCTGAAGACCGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCT 1753
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4415 AGGAGCCCTTCAAGAACCTGAAGACCGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCC 4474

Qy   1754 ACGACGTGAAGCAGTCTGACCGAGCCCATGGAAGGAGACCTGGACATCGTGATCT 1813
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4475 ACGACGTGAAGCAGTCTGACCGAGCCCATGGAAGGAGACCTGGACATCGTGATCT 4534

Qy   1814 GGGGCAAGAACCCCCAAGTTCCGCCTGCCCATCCGACCTGTGGACCTGGTGGA 1873
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4535 GGGGCAAGAACCCCCAAGTTCCGCCTGCCCATCCGACCTGTGGACCTGGTGGA 4594

Qy   1874 CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGG 1933
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4595 CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGG 4654

Qy   1934 TGAAGCTGTGGTACCAGCTGGAGAAGGCCCATCATCGGCGGCGGCTACGTGG 1993
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4655 TGAAGCTGTGGTACCAGCTGGAGAAGGCCCATCATCGGCGGCGGCTACGTGG 4714

Qy   1994 ACGGCGCGCCAACCGCGAGAACAAGATCGGCAAGAGCGGCGGGCC 2053
         |||||||||||||||||||||||||||||||||||||||||||||
Db   4715 ACGGCGCGCCAACCGCGAGAACAAGATCGGCAAGAGCGGCGGGGCC 4774

Qy   2054 GCCAGAAGATCGTGAGGCTGACCGAGACCGAGCTGCAGGCCATCC 2113
         |||||||||||||||||||||||||||||||||||||||||||||
Db   4775 GGCAGAAGATCGTGAGGCTGACCGAGACCGAGCTGCAGGCCATCC 4834

Qy   2114 AGCTGGCCCTGACGCGGCAGCGCGAGGTGAACATCGTGACGACGCCAGTAGCGCC 2173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4835 AGCTGGCCCTGACGCGAGGACAGCGCGAGGTGAACATCGTGACGACGCCAGTAGCGCCCC 4894

Qy   2174 TGGGCATCATCCGACCTGTGCCCCGGGCTGGAAGGGCAGCCCCAGCATCATCG 2233
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4895 TGGGCATCATCCGACCTGTGCCCCGGGCTGGAAGGGCAGCCCCAGCATCATCG 4954

```
Qy   2234  AGCAGTCGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG 2293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4955  AGCAGTCGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG 5014

Qy   2294  GCGGGCAACGACGAGATCGACAAGCTGGTGAGCAAGGGCATCGCCAAGGTGGCTGTTCCTGG 2353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5015  GCGGGCAACGACGAGATCGACAAGCTGGTGAGCAAGGGCATCGCCAAGGTGGCTGTTCCTGG 5074

Qy   2354  ACGGCATCGATGGCGGCAGACTCGTGATCGATCTACCAGTACGTGGACAGACGTGTGGGCAGCG 2413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5075  ACGGCATCGATGGCGGCAGACTCGTGATCGATCTACCAGTACAATGGACAGACGTGTGGGCAGCG 5134

Qy   2414  GCGGGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
           ||||||||||||||||||||||||||||||||||||||||||||
Db   5135  GCGGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 5178
```

RESULT 15
ADC13234
ID   ADC13234 standard; DNA; 3531 BP.
XX
AC   ADC13234;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagPolmut_C SEQ ID NO 13.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 10; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramuscosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3531 BP; 802 A; 1210 C; 1096 G; 423 T; 0 U; 0 Other;

Query Match   96.8%; Score 2383.6; DB 9; Length 3531;
Best Local Similarity 98.8%; Pred. No.1.9e-287;
Matches 2414; Conservative   0; Mismatches 24; Indels 6; Gaps 1;
```

```
Qy    14  TGGCCGAGCCATGAGCCATGAGCCAGGCCAACATCTGCTGATGCAGCGCGCAGCCGCCAACTTCA  73
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1088  TGGCCGAGCGGATGAGCCATGAGCCAGGCCAACACCAGCGTGATGCAGCGCGAAGAGCAACTTTAAAA 1147

Qy    74  AGGGCCCCAAGGCGGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCCA 133
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1148  AGGGCCCCAAGGCGGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCCGCCA 1207

Qy   134  ACTGCCGCGCCCCGCCCCGCAAGAAGAAGGGCTGCTGGAAGTGCGGCAAGTGCGGCCACCAGATGA 193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208  ACTGCCGCGCCCCGCCCCGCAAGAAGGGCTGCTGGAAGTGCGGAAGTGCGGCAAGTGCGGCCACCAGATGA 1267

Qy   194  AGGACTGCACCGAGCGGCGCCAGGCCAACTTCTTCCGGGACGGCGCCAACTTCTTCCGGGAC 253
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268  AGGACTGCACCGAGCGGCGCCAGGCCAACTTCTTCCGGGACGGCGCCAACTTCTTCCCAGGGCA 1327

Qy   254  AGGCCCGCGAGTTCCCGCAGCGAGCAGAACCGCGCCAACAGCCCGCAGCCCGCGAGCTGC 313
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328  AGGCCCGCGAGTTCCCGCAGCGAGCAGAACCGCGCCAACAGCCCGCAGCCCGCGAGCTGC 1387

Qy   314  AGGTGCGGCGGCGCAACCCCGCAGCGAGGCCGAGGCCAGGGCCACCCTGAACT 373
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388  AGGTCGCGGCGCGCAACCCCGCAGCGAGGCCGAGGCCAGGGCCACCCTGAACT 1447

Qy   374  TCCCCCAGATCACCCTGTGGCAGCGCGCCCCTGTGAGCATCAAGGTGGCGGCCAGATCA 433
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1448  TCCCCCAGATCACCCTGTGGCAGCGCGCCCCTGTGAGCATCAAGGTGGCGGCCAGATCA 1507

Qy   434  AGGAGGGCCCTGCTGGACACCGGCGGCGGCGCGACAACCGTGCTGACACCGTGGGCTGCCGCCG 493
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1508  AGGAGGGCCCTGCTGGACACCGGCGGCGCGACAACCGTGCTGTGGGAGGATGAGCCTGCCGCCG 1567

Qy   494  GCAAGTGGAAGCCCAAGATGATCGGCGCCTTCATCAAGGTGCGCCAGTACG 553
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1568  GCAAGTGGAAGCCCAAGATGATCGGCGCCTTCATCAAGGTGCGCCAGTACG 1627

Qy   554  ACCAGATCCTGATGAGCCTGCCCGTGAAGCCCATCGGCACCGTGCTGATCGGCCCCA 613
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1628  ACCAGATCCTGATGAGCCTGCCCGTGAAGCCCATCGGCACCGTGCTGATCGGCGCCCCA 1687

Qy   614  CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1688  CCCCCGTGAACATCATCGGCCGCAACATCTGCACCCAGCTGGGCTGCACCCTGAACTTCC 1747

Qy   674  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGGCCGGCCCCAGG 733
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1748  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGGCCGGCCCCAGG 1807

Qy   734  TGAAGCAGTGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGACTTCAGGGAGGAGA 793
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1808  TGAAGCAGTGCGCCCTGACCGAGGAGATCAAGGCCCTGACCGACTTCGGAGGAGA 1867

Qy   794  TGGAGAAGGAGGGGCAAGATCACCAAGATCGGGCCCCGAGAACCCCTACACACCCCGTGT 853
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1868  TGGAGAAGGAGGGGCAAGATCACCAAGATCGGGCCCCGAGAACCCCTACACAACACCCCGTGT 1927

Qy   854  TCGCCATCAAGAAGGAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGGAGCTGA 913
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1928  TCGCCATCAAGAAGGAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGGAGCTGA 1987

Qy   914  ACAAGCGCCACCCAGGACTTCTGGGAGGTGCAGTCGGCATCCCCACCCCCACCCGGCCTGA 973
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1988  ACAAGCGCCACCCAGGACTTCTGGGAGGTGCAGTCGGCATCCCCACCCCCACCCGGCCTGA 2047

Qy   974  AGAAGAAGAGAGCGGCTGCCGTGCTGGACGTGGCGACGTGGCTACTTCAGCGTGCCCTG 1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2048  AGAAGAAGAGAGCGGCTGCCGTGCTGGACGTGGCGACGTGGCTACTTCAGCGTGCCCTG 2107

Qy  1034  ACGAGGACTTCCGCAAGTACACCACGGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 1093
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2108  ACGAGGACTTCCGCAAGTACACCACGGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 2167
```

```
Qy   1094  GCATCCGGCTACCAGTACACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC  1153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2168  GCATCCGGCTACCAGTACACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC  2227

Qy   1154  AGAGCAGCATGACCAAGAGATCCTGGAGCCCTTCCGCGCCGCAACCCCGAGATCGTGATCT  1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2228  AGAGCAGCATGACCAAGAGATCCTGGAGCCCTTCCGCGCCGCAACCCCGAGATCGTGATCT  2287

Qy   1214  ACCAGGCCCCCTGTACGTGGGCAGCGACCTGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCG  1273
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2288  ACCAGGCCCCCTGTACGTGGGCAGCGACCTGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCG  2347

Qy   1274  AGGAGCTGCGCAAGCACCTGCTGCGGTGCGGCTTCACCACCCCCGACAAGAGAGCACCAGA  1333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2348  AGGAGCTGCGCAAGCACCTGCTGCGGTGCGGCTTCACCACCCCCGACAAGAAGAGCACCAGA  2407

Qy   1334  AGGAGCCCCCTTCCTGTGGATGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGC  1393
           ||||||||||||||||       |||||||||||||||||||||||||||||||||||
Db   2408  AGGAGCCCCCTTCCTGCCCAT-----CGAGCTGCACCCCGACAAGTGGACCGTGCAGC  2461

Qy   1394  CCATCGAGCTGCCCGAGAACGAGAGCTGGACCCTGAACGACATCTCAGAAGCTGTGGGCA  1453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2462  CCATCGAGCTGCCCGAGAACGAGAGCTGGACCCTGAACGACATCTCAGAAGCTGTGGGCA  2521

Qy   1454  AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC  1513
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2522  AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC  2581

Qy   1514  TGCGCGGCGCCAAGGCCTGACCGAGCATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGC  1573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2582  TGCGCGGCGCCAAGGCCCTGACCGAGCATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGC  2641

Qy   1574  TGGCCGAGAACCGGCAGATCCTGCGCGAGATCCTGCGGCAGCCCGTGCACGCGTGTACTACGACCCCCAGCA  1633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2642  TGGCCGAGAACCGGCAGATCCTGCGGCGAGATCCTGCGGCAGCCCGTGCACGCGTGTACTACGACCCCCAGCA  2701

Qy   1634  AGGACCTGGTGGCCGAGAATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACC  1693
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2702  AGGACCTGGTGGCCGAGAATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACC  2761

Qy   1694  AGGAGCCCTTCAAGAACCGGCAAGTACGCCAAGAATGCGCACCGGCCCACCACCA  1753
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2762  AGGAGCCCTTCAAGAACCGGCAAGTACGCCAAGAATGCGCACCGCCCCACCACCA  2821

Qy   1754  ACGACGTGAAGCAGCTGCAGCCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCT  1813
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2822  ACGACGTGAAGCAGCTGCAGCCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCT  2881

Qy   1814  GGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGA  1873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2882  GGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGA  2941

Qy   1874  CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCCCCTGG  1933
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2942  CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCCCCCTGG  3001

Qy   1934  TGAAGCTGTGTACAGCCGGCAAGGCCGCAAGACCAGAGACCGGAGCTGCAGGCTGACGTGG  1993
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3002  TGAAGCTGTGTGGTACCAGCCGGCAAGGCCGCAAGACCAGAGACCGGAGCTGCAGGCCATCC  3061

Qy   1994  ACGGCCGCCAACCGCCGCCAAGGATCGGCACACCGCCAAGATCGGCACCGGTGACCGCAGCGTACGCCGCC  2053
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3062  ACGGCCGCCAACCGCCGCCAAGGATCGGCACACCGCCAAGATCGGCACCGGTGACCGCAGCGTACGCCGCC  3121

Qy   2054  GGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGACCGGAGCTGCAGGCGCATCC  2113
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3122  GGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGACCGGAGCTGCAGGCCATCC  3181

Qy   2114  AGCTGGCCCTGACCGGACGAGCGCGAGCGCGCCAGGTGAAACATCGTGACCGACACCAGTACGCCC  2173
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3182  AGCTGGCCCTGACCGGACGAGCGCGAGCGCGCCAGGTGAACATCGTGACCGACACCAGTACGCCC  3241

Qy   2174  TGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCG  2233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3242  TGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCG  3301

Qy   2234  AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG  2293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3302  AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG  3361

Qy   2294  GCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGCCATCGCCGCAAGGTGCTGTTCCTGG  2353
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3362  GCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGCATCCGCCAAGGTGCTGTTCCTGG  3421

Qy   2354  ACGGCATCGATGGCGGCATCGTGATCGATCTACCAGTACATGACGACGACCTGTACGTGGGCAGCG  2413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3422  ACGGCATCGATGGCGGCATCGTGATCGATCTACCAGTACATGACGACGACCTGTACGTGGGCAGCG  3481

Qy   2414  GCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2457
           ||||||||||||||||||||||||||||||||||||||||||||
Db   3482  GCGGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT  3525
```

```
                     GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    April 10, 2004, 06:48:32 ; Search time 118.333 Seconds
                              (without alignments)
                              11550.795 Million cell updates/sec

Title:      US-09-610-313-31
Perfect score:  2463
Sequence:   1 gtcgacggccaccatggccga.........gggctagcaccggtgaattc 2463

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:   682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:   1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Issued_Patents_NA:*
            1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
            2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
            3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
            4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
            5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
            6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                           SUMMARIES
```

Result No.	Score	% Query Match	Query Length	DB	ID	Description
1	2046	83.1	2306	4	US-09-475-515-82	Sequence 82, Appli
2	2025.2	82.2	2312	4	US-09-475-515-84	Sequence 84, Appli
3	2019.2	82.0	2300	4	US-09-475-515-83	Sequence 83, Appli
4	1942.6	78.9	4319	4	US-09-475-515-6	Sequence 6, Appli
5	1878.8	76.3	2305	4	US-09-475-515-80	Sequence 80, Appli
6	1852	75.2	2299	4	US-09-475-515-81	Sequence 81, Appli
7	1651.8	67.1	4307	4	US-09-552-950-2	Sequence 2, Appli
8	1624.6	66.0	9772	4	US-09-552-950-5	Sequence 5, Appli
9	1566	63.6	8366	4	US-09-872-733A-6	Sequence 6, Appli
10	1530.6	62.1	4338	4	US-09-872-733A-1	Sequence 1, Appli
11	1203.6	48.8	9010	4	US-09-184-418C-8	Sequence 8, Appli
12	1172.6	47.6	8972	4	US-09-184-418C-9	Sequence 9, Appli
13	1165.4	47.3	8959	4	US-09-184-418C-11	Sequence 11, Appli
14	1163.8	47.3	2467	4	US-09-872-733A-3	Sequence 3, Appli
15	1142	46.4	8992	4	US-09-184-418C-4	Sequence 4, Appli
16	1106.8	44.9	2601	3	US-09-117-217-7	Sequence 7, Appli
17	1106.8	44.9	2601	3	US-09-117-217-9	Sequence 9, Appli
18	1106.8	44.9	2601	3	US-09-117-217-11	Sequence 11, Appli
19	1106.8	44.9	2601	3	US-09-117-217-13	Sequence 13, Appli
20	1106.8	44.9	2601	4	US-09-735-487-7	Sequence 7, Appli
21	1106.8	44.9	2601	4	US-09-735-487-9	Sequence 9, Appli
22	1106.8	44.9	2601	4	US-09-735-487-11	Sequence 11, Appli
23	1106.8	44.9	2601	4	US-09-735-487-13	Sequence 13, Appli
24	1106.8	44.9	4307	4	US-09-552-950-1	Sequence 1, Appli
25	1106.2	44.9	9719	4	US-09-700-304-1	Sequence 1, Appli
26	1105.2	44.9	9050	4	US-09-184-418C-7	Sequence 7, Appli
27	1100.4	44.7	7399	2	US-08-418-848A-9	Sequence 9, Appli
28	1100.4	44.7	9709	2	US-08-188-583-5	Sequence 5, Appli
29	1100.4	44.7	9709	3	US-08-388-353-1	Sequence 1, Appli
30	1100.4	44.7	9709	3	US-08-488-551B-1	Sequence 1, Appli
31	1100.4	44.7	9709	4	US-09-309-572-15	Sequence 15, Appl
32	1100.4	44.7	9709	4	US-09-718-096-15	Sequence 15, Appl
33	1100.4	44.7	12479	4	US-09-318-138-13	Sequence 13, Appl
34	1100.4	44.7	12494	3	US-08-935-312-13	Sequence 13, Appl
35	1100.4	44.7	12494	3	US-08-848-760B-33	Sequence 33, Appl
36	1100.4	44.7	15581	3	US-08-646-538-35	Sequence 35, Appl
37	1100.4	44.7	15581	3	US-09-503-222-35	Sequence 35, Appl
38	1099.2	44.6	3000	4	US-09-184-418C-74	Sequence 74, Appl
39	1095.8	44.5	8968	4	US-09-184-418C-1	Sequence 1, Appli
40	1090.8	44.3	4313	5	US-09-475-515-3	Sequence 3, Appli
41	1090.8	44.3	9737	2	US-08-944-449-7	Sequence 7, Appli
42	1090.8	44.3	9737	4	US-09-353-362-7	Sequence 7, Appli
43	1082.8	44.0	8954	4	US-09-184-418C-6	Sequence 6, Appli
44	1077.6	43.8	3017	4	US-09-184-418C-83	Sequence 83, Appl
45	1075.4	43.7	3011	4	US-09-184-418C-101	Sequence 101, App

```
                          ALIGNMENTS

RESULT 1
US-09-475-515-82
; Sequence 82, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
US-09-475-515-82
;  SEQ ID NO 82
;  LENGTH: 2306
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: FS(-).protmod.RTopt.YM
US-09-475-515-82

Query Match             83.1%;  Score 2046;  DB 4;  Length 2306;
Best Local Similarity   93.6%;  Pred. No. 0;
Matches  2159;  Conservative  0;  Mismatches  135;  Indels  12;  Gaps  2;

Qy      170 GCGGCAAGGAGGAGGCCACCAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCCGCG 229
                ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        1 GCGGCCCGCGAAGGACGACCACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTCTTCCGCG 60

Qy      230 AGGACCTGGCCTTCCCCCAGGCGAGCGAGTTCCCGGAGCAGAACCGCGCCA 289
                ||||||||||||| |||||| |||||||||||||||| ||||||||||||
Db       61 AGGACCTGGCCTTCGCAGGGCAAGGCCGCGAGTTCAGCAGCAGCAGACCCGCGCCA 120

Qy      290 ACAGCCCCACCAGCCGCGAGCTGCAGGTGCGCGGGG------ACAACCCCGCCAGCGAGG 343
                ||||||||||||||||||| |||||| ||||  |||   |||| ||||||||||||||||
Db      121 ACAGCCCCACCACCCGCGAGCTGCAGGTGTGGGGCGGCGGAGAACAACAGCCTGAGCGAGG 180

Qy      344 CCGGCGCCGAGCGCCAGGGCCACCCTG------AACTTCCCCAGATCACCCTGTGCAGC 397
                |||||||||||||||||||||||||||    |||||||||||||||||||||||||||||
Db      181 CCGGCGCCGACCCGCCAGGGCACCGTGAGCTCAACTTCCCCAGATCACCCTGTGCAGC 240
```

```
Db 1321 ACGCCCGGCATCAAGGTGAAGCTGCAGCTGTGCCAAGCTGCTGCCGGCGGCCACCGGCCCCTGACCG 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1538 ACATCGTGCCCCTGACCGGAGGAGGCCAGCTGTGCCAAGCTGCTGCGGCGGCCACCGGCCCCGAGATCCTGC 1597

Db 1381 AGGTGATCCCCCTGACCGGAGGAGGCCGAGCTGGGACCTGGGCCGAGACCGGCCGAGACCGGCGAGATCCGCA 1440
Qy 1598 GCGAGCCCGTGCACGGCGTGCTACGGACCTCGTGGCCGAGATCCAGA 1657

Db 1441 AGGAGCCGTGCACGAGGTGCTACTACGAACCCCAGCAAGAAGGACCTGTGGCCGAGATCCAGA 1500
Qy 1658 AGCAGGGCCACGACCAGTGACCTACCAGATCTACCAGAGAGCCCTTCAGCAAACCTGAAGA 1717

Db 1501 AGCAGGGCCAGGGCCCAGTGGCCACTGGACCTACCAGATCTACCAGGAGACCCTTCAAGCAAACCTGAAGA 1560
Qy 1718 CCGGCAAGTACGACCCAAGATCGGCCACCGGCACCCCACAACGCGTGAAGCAGCTGACCGAGG 1777

Db 1561 CCGGCAAGTACGCCCGCATGCGGCGCGCCGCCACCAACGACGTGAAGCAGCTGACCGAGG 1620
Qy 1778 CCGTGCAGAAGATCGCCCAAGGTGAGCATCGTGGGCGCAGGGCAGGAGATCGTGCAGAAGAGACCCCAAGTTCCGCC 1837

Db 1621 CCGTGCAGAAGATCGAGCCAGGAGATCGGGGCAGGCAGGCAGGGGCAAGAGATCGTGGGTGGCCAGGTTCAAGC 1680
Qy 1838 TGCCCATCCAGAAGGAGACCTGGGACACCTGGTGACGACTGCAGGCCACCTGAGA 1897

Db 1681 TGCCCATCCAGAAGGAGACCTGGGAGGGCCTGGTGGAACGTGGTGTGTACCAGCTGAGA 1740
Qy 1898 TCCCCGAGTGGAGTTCGTGAACACCCCCCTGTGAAGCTGTGTGTACCAGTGTGAGA 1957

Db 1741 TCCCCGAGTGGAGTTCGTGAACACCCCCCTGTGAAGCTGTGTGTACCAGTGTGAGA 1800
Qy 1958 AGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGACGGCGCGCGCCAACCGCGAGACCG 2017

Db 1801 AGGAGCCCATCGTCGGGGCCGAGACCCTTCTACGTGACGGGGCGCGCCGGCCAACCCGAGACCA 1860
Qy 2018 AGATCGCGCAAGGCCGGCTACGTGACGAGGCCAAGGATCGTGAGCGTGACCG 2077

Db 1861 AGCTGGGCAAGGCCGGCTACGTGACGAGGCCAGGGTTGGTGAGCATCGCCG 1920
Qy 2078 AGACCACCAACCAGAAGACGAGCTGCAGGACTGCAGAGCAGCCGGCA 2137

Db 1921 ACACCACCAACCAGAAGACGAGCTGCAGGACTGCAGAGCAGCCGGCC 1980
Qy 2138 GCGAGGTGAACATCGTGACCGACCAGCCAGTGAGCATCATCCAGGCCCAGCCCG 2197

Db 1981 TGGAGGTGAACATCGTGACCGACCAGCCAGTGAGCATCATCCAGGCCCCAGCCCG 2040
Qy 2198 ACAAGAGGCGAGAGCGAGCTGGTGAACCAGACAGCTGAAGAGAGGAGAAGG 2257

Db 2041 ACAAGAGGCGAGAGCGAGCTGGTGAACCAGACAGCTGAAGAGAGGAGAAGG 2100
Qy 2258 TGTACCTGAGCTGGGTGCCCGCCACAAGGGCATCGCAACGAGCAGATCGACAAGC 2317

Db 2101 TGTACCTGGCCTGGGTGCCCGCCACAAGGGCATCGGGCGCAACGAGCAGGTGGACAAGC 2160
Qy 2318 TGGTGAGCAAGGCGCATCCGGACGTGCTGTTCCTGGCAGCGGCGGATGCGGCCATCGTGA 2377

Db 2161 TGGTGAGCGCGCATCCGGACGTGCTGTTCCTGGCAGCGGCGGATGCGGCCATCGTGA 2220
Qy 2378 TCTACCAGTACATGAACGACGCCGCGCGCCGGCCGCGCCTAGGATCGATTAAAAGC 2437

Db 2221 TCTACCAGTACATGAACGACGCCGCGCGCCGGCCGCGCCTAGGATCGATTAAAAGC 2280

Qy 2438 TTCCCGGGGCTAGCACCGGTGAATTC 2463
Db 2281 TTCCCGGGGCTAGCACCGGTGAATTC 2306
```

RESULT 2
US-09-475-515-84
; Sequence 84, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:

```
Qy 398 GCCCCTGTGAGCATCAAGGTGGCGGCCAGATCAAGCAGGAGGCCCTGCTGGACACCCGCG 457
          ||||||||||||||||||||||||||||||||||||||||
Db 241 GCCCCTGTGTGACCATCAGGATCAGGATCGGCGGCCAGCTCAAGGAGGCGCGTGCTCCGACACCGGCG 300

Qy 458 CCGACGACACCGTGCTGCTGGAGATGAGCGATGAGCCTGCCCGGCAAGTTGGAAGCCCAAGATGATCG 517
Db 301 CCGACGACACCGTGCTGCGTGGAGGATGAGCGATGAACCTGCCCGGCGAGTGGAAGCCCAAGATGATCG 360

Qy 518 GCGGCATCGGGCGGCGGTTCATCAAGTGCGCGCAGTACGACCAGATCCTGATCGAGATCTGCG 577
Db 361 GCGGGATCGGGGGCGGCTTCATCAAGTGCGCGCAGTACGACCAGATCCCCGTGGAGATCTGCG 420

Qy 578 GCAAGAAGGCCACGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGAACATCATCGGCGCCGCA 637
Db 421 GCCCAAGGCCATCGGCACCGTGCTGGTGGTGGGCCCCACCCCCGTGAACATCATCGGCCCCA 480

Qy 638 ACATGCTGACCCAGCTGCGGCTGACCCTGAACTTCCCCATCAGCCCATGCGGACCGTGC 697
Db 481 ACCTGCTGACCCAGATCGGCTGACCCTGAACTTCCCCATCAGCCCATGCGGACGACGGGTGC 540

Qy 698 CCGTGAAGCTGAAGCCCGGCATGACGGACGCTGGAGCCAGTGGCCCCTGACCGAGG 757
Db 541 CCGTGAAGCTGAAGCCGGGCGGATGACGCGGCCCAGTGTCAAGCCAGTGGCCCCTGACCGAGG 600

Qy 758 AGAGAATCAAGGCCCTGACCGCCGACACACAACACCCCGTGTTCGCCATCAAGAGAAGGACAGCA 817
Db 601 AGAGAATCAAGGGCCCCTGGTGGAGATCTGCACCCAGACACACCCCCGTGTTGTCCGCCATCAAGAGAAGGACAGCA 660

Qy 818 AGATCGGCCCCGAGAACCCCTACACACACCCCCGTGTTCGCCATCAAGAGAAGGACAGCA 877
Db 661 AGATCGGCCCCGAGAACCCCTACACACACCCCCGTGTTGTCCGCCATCAAGAGAAGGACAGCA 720

Qy 878 CCAAGTGGGCCAAGGTGGTGGACTTCCGGGAGCTGAACAAGCGCACCCCAGGACTTCTGGG 937
Db 721 CCAAGTGGGCCAAGGTGGTGGACTTCCGGGAGCTGAACAAGCGCACCCCAGGACTTCTGGG 780

Qy 938 AGGTGCAGCTGGGCATCCCCCACCCCCGGCCTGAAGCGCTACACCAGTGACCGTGC 997
Db 781 AGGTGCAGCTGGGCATCCCCCACCCCCGGCCTGAAGCGCTACACCAGTGACCGTGC 840

Qy 998 TGGACCTGTGGGGCGACGCGCTCTTCAGCGTGTGCCCTGGACGTGCCCGCCTACCGAGGACTTCCGCGAGGACTACACCG 1057
Db 841 TGGACCTGTGGGGCGACGCGCTCTTCAGCGTGTGCCCTGGACGTGCCCGCCTACCGAGGACTACACCG 900

Qy 1058 CCTTCACCACCCCAGCATCAACAACGAGACCCCCGGCATCTTCGACGAGCATGACCAAGAGTCTGC 1117
Db 901 CCTTCACCACCCCAGCATCAACAACGAGACCCCCGGCATCTTCGACGAGCATGACCAACGTGC 960

Qy 1118 TGCCCCAGGGGCTGGAAGGGCAGCAGCCCCAGCATCTTCAGAGCAGCATGACCAAGAATCCTGC 1177
Db 961 TGCCCCAGGGCTGGAAGGGCAGCAGCCCCAGCATCTTCAGAGCAGCATGACCAAGAATCCTGC 1020

Qy 1178 AGCCCTTCCGGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGGCA 1237
Db 1021 AGCCCTTCCGGCCCGCAACAGCCCCGAGAAACATCGTGATCTACCAGGCCCCCCTGTACGTGGGGCA 1080

Qy 1238 GCGACCTGGAGATCGGCCGACAAGTGACCCGCCGCAGCAGATCGGCCGAGCTGGAGGCGACCACCTGCTGC 1297
Db 1081 GCGACCTGGAGATCGGCCGACCACCGGCCAGATCGGCCGAGCTGGAGGCGACCACCTGCTGC 1140

Qy 1298 GCTGGGGGCTTCACCACCCCCGACACGTTGCAAGCTGAACTGGCAGCTGGTGCCAAGATCT 1357
Db 1141 GCTGGGGGCTTCACCACCCCCGACACGTTGGTGGCAAGCTGAACTGGCAGCTGGTGCCCAGCCAGATCT 1200

Qy 1358 GCTACGAGCTGACCCCGACAAGTGACAAGTGACCAGCCCATCGAGCTGCCCGAGAAGGAGA 1417
Db 1201 GCTACGAGCTGACCCCGACAAGTGACAAGTGACCAGCCCATCATGCTGCCCGAGAAGGACA 1260

Qy 1418 GCTGGACCGTGACAGCTCGAGACATCCAGAAGCTGGTGCGCAAGCTGAACTGGGCCCCAGATCT 1477
Db 1261 GCTGGACCGTGACCGAACGACATCCAGAAGCTGGTGGCGCAAGCTGAACTGGGCCCCAGATCT 1320

Qy 1478 ACCCCGGCATCAAGGTGCGCCAGCTGCGGCGCGCCGCCAGCTGCTGCGCGGCCCAAGGGCCCCTGACCG 1537
```

RESULT 2
US-09-475-515-84 2463
; TTCCCGGGGCTAGCACCGGTGAATTC 2306

```
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 84
; LENGTH: 2312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(-).protmod.RTopt(+)
US-09-475-515-84

Query Match      82.2%;   Score 2025.2;   DB 4;   Length 2312;
Best Local Similarity   93.3%;   Pred. No. 0;
Matches  2156;   Conservative   0;   Mismatches  138;   Indels  18;   Gaps   3;
```

```
Db   1741 ||||||||||||||||||||||||||||||||||||||||||||||||||
          CCTGATCCCGAGTGGGAGTTCGTGACACACCCCCCCTGGTGAAGCTGTGGTACCAGC 1800

Qy   1952 TGGAGAAGGAGCCACCATCATCGGCGCCAAGAGAGTTCTACGTGACGACGACCAACCGCG 2011
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TGGAGAAGGAGCCACCATCATCGGCGCCAAGAGAGTTCTACGTGACGACGACCAACCGCG 1860

Qy   2012 AGACCAAGATCGGCAAGGCCGGCTACGTGACGACCGGGCCGGCAGAAGATCGTGCAGCC 2071
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 AGACCAAGATCGGCAAGGCCGGCTACGTGACGACCGGGCCGGCAGAAGGTGGTGAGCA 1920

Qy   2072 TGACCGAGACCACCACCAACCAGAAGACCGAGCTGAGCCATCCAGCTGACCTGGCCCTGCCAGGACA 2131
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TCGCCGACACCACCAACCAGAAGACCGAGCTGCCAGGCCATCCACCTGGCCCTGCCAGGACA 1980

Qy   2132 GCGGCGGAGGTGAACATCTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC 2191
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GCGGCCTGGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC 2040

Qy   2192 AGCCCGACAAGAGAGGCGGAGGCGAGTCGGTGAACCAGATCATCGAACAGCTGATCAAGAGAGG 2251
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AGCCCGACAAGAGCGGCGGAGGCGAGTCGGTGAACCAGATCATCGAACAGCTGATCAAGAAGG 2100

Qy   2252 AGAAGGTGTACCTGAGCTGGGTGCCCCACAAGGGCATCGGCGGCAACGAGCAGATCG 2311
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 AGAAGGTGTACCTGGCCTGGGTGCCCCACAAGGGCATCGGCGGCAACGAGCAGGTGG 2160

Qy   2312 ACAAGCTGGTGAGCAAGGCCATCCGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCA 2371
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 ACAAGCTGGTGAGCGGCGCCATCCGCAAGGTGCTGTTCCTGAACGGCATCGATGCGGCA 2220

Qy   2372 TCGTGATCTCACCAGTACGATGACGACCTGTACGTGGGCAGCGGCGGGCCGCGGGATGATT 2431
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TCGTGATCTCACCAGTACGATGACGACCTGTACGTGGGCAGCGGCGGGCCGCGGGATGATT 2280

Qy   2432 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2463
          ||||||||||||||||||||||||||||||||
Db   2281 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2312
```

```
Qy   170 GCGGCAAGGAGGAGGGCCACCAGATGAGGACTGCACCGAGCGCGCCAGGGCGCCAACTTCTTCCGCG 229
         ||| ||| ||| ||||||||||||||||||||||||||||||||||||
Db     1 GCGGCCGCGAAGGACACCAAATGAAAGATTCACTGAGAGAAAGGCTAATTTCTTCCGCG 60

Qy   230 AGGACCTGGCTTCCCCAGGCAAGGCCGCGAGTTCCCAGCGAGCAGAACCGGCCA 289
         |||||||||||||||||||||||||||||||||||||
Db    61 AGGACCTGGCCTTCTGCAGGCACCGGGCCCGGCAGTTCAGCAGCAGCCCGCGCCA 120

Qy   290 ACAGCCCCACCACCACCGCCGCCGAGTGCGGGG------ACAACCCCGCAGCGAGG 343
         ||||||||||||||||||||||||||
Db   121 ACAGCCCACCACCACCGCCGCCGAGTGTGGGGCGGCGACAACAGCCTGAGCGAGG 180

Qy   344 CCGGCGCCGAGCGCCAGGGCACCCTG------AACTTCCCCAGATCACCTGTGCAGC 397
         |||||||||||||||
Db   181 CCGCCGCCGACCGCCCGCCAGGCACCGTGAGCTTCAACTTCCCCAGATCACCCGTGTGCAGC 240

Qy   398 GCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGGCCCTGCTGGACACCGGCA 457
         |||||||||||||||||||||||||||||||||||||||
Db   241 GCCCCCTGGTGACCATCAAGGTGGGGCGGCCGGCAGTCAAGGAGGGCGCTGCTCGACACCGGGG 300

Qy   458 CCGAGGACACCGTGCTGACGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCG 517
         |||||||||||||||||||||||||||||||||||||
Db   301 CCGAGGACACCGTGCTGACGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCG 360

Qy   518 GCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACGACGACCAGATCCTGATCGAGATCTGCG 577
         ||||||||||||||||||||||||||||||||||||||||
Db   361 GCGGGATCGGCGGCTTCATCAAGGTGCGGCCAGTACGACGACCAGATCCCGTGGAGATCTGCG 420

Qy   578 GCAAGAAGCCATCGGCACCGTGCTGATCGGCCCCCGTGACATCATCGGCCGCGCA 637
         ||||||||||||||||||||||||||||||||||||
Db   421 GCCACAAGGCCATCGGCACCGTGCTGTGGTGGCAAGTGCACCCCGTGAACATCATCGGGCGCCA 480

Qy   638 ACATGCTGACCCAGCTGGGCTGCACCGCCCCATCAGCCCCATCGAGACCTGC 697
         |||||||||||||||||||||||||||||||||||||
Db   481 ACCTGCTGACCCAGCTGGGCTGCACCGCCCCATCAGCCCCATCGAGACCGGTGC 540

Qy   698 CCGTGAAGCTGTGAAGCCCGCATGGACGGCCCGGAGGCAGCCAAGAGCGGTCACCGAGG 757
         ||||||||||||||||||||||||||||||||||||||
Db   541 CCGTGAAGCTGTGAAGCCCGGGATGGACGGCCCGGGAGGCAGCAGTGCCCCTGACCGAGG 600

Qy   758 AGAAGATCAAGGCCCTGACCGCCATCTGCGAGGATGGAGAAGGAGAGGCAAGATCACCA 817
         |||||||||||||||||||||||||||||||||||||
Db   601 AGAAGATCAAGGGCCCTGGTGGAGATCTGCACCGAGGATCTGGAGAAGGAGAAGGCAAGATCAGCA 660

Qy   818 AGATCGGCCCCGAGAACCCCTACACACACCCCGGTGTTCGCCATCAAGAAGAAGGAACAGCA 877
         |||||||||||||||||||||||||||||||||||||
Db   661 AGATCGGCCCCGAGAACCCCTACACACACACCCCGTGTTCGCCATCAAGAAGAAGAAGGAACAGCA 720

Qy   878 CCAAGTGGGCCAAGCTGGTGGACTTCCGGCAGCTGAACAAGGCCACCCAGGACTTCTGG 937
         |||||||||||||||||||||||||||||||||||||
Db   721 CCAAGTGGGCCAAGCTGGTGGACTTCCGGCAGCTGAACAAGGCCACCCAGGACTTCTGG 780

Qy   938 AGGTGCAGTGGGCATCCCCGGCCTCCCACCCCGCCGCCGGCGTGACCGTGCAGATCCTGC 997
         |||||||||||||||||||||||||||||||||||||
Db   781 AGGTGCAGTGGGCATCCCCGGCCTGGGCATCCCCGCCGCCGGCGTGACCGTGCAGATCCTGC 840

Qy   998 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGTGGCCATCTTCAGCAAGTACACCG 1057
         |||||||||||||||||||||||||||||||||||||
Db   841 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACAAGTTCAGCAAGTACACCG 900

Qy  1058 CCTTCACCCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 1117
         |||||||||||||||||||||||||||||||||||||
Db   901 CCTTCACCCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC 960

Qy  1118 TGCCCCAGGGCTGAAGGGCAGCCCCAGATCTTCCAGCAGCCCCAAGATCCTG 1177
         |||||||||||||||||||||||||||||||||||||
Db   961 TGCCCCAGGGCTGAAGGGCAGCCCCGCCATCTTCCAGCAGCCCCAAGATCCTGG 1020

Qy  1178 AGCCCTTCCGCGCCCCGCCAACCCCGAGATCGTGATCTTGTGATCTGATCGATCGACGTGC 1237
         |||||||||||||||||||||||||||||||||||||
Db  1021 AGCCCTTCCGCGCCAAGCAGCAGAGACCCCGACATCGTGATCGTGATCGACGTGC 1080
```

```
RESULT 3
US-09-475-515-83
; Sequence 83, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 83
;  LENGTH: 2300
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: FS(-).protmod.RTopt.YMWM
US-09-475-515-83

Query Match             82.0%;  Score 2019.2;  DB 4;  Length 2300;
Best Local Similarity   93.2%;  Pred. No. 0;
Matches  2150;  Conservative  0;  Mismatches  138;  Indels  18;  Gaps  3;
```

us-09-610-313-31

```
Qy  1238 GCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1297
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 GCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1140

Qy  1298 GCTGGGCTTCACCACCCCCGACAGAGAGCACCAGAGGAGCCCCCTTCCTGTGTGATGG    1357
         ||||||||||||||||||||                    ||||||||||||||||
Db  1141 GCTGGGGCTTCACCACCCCCGACACCCCGACAAGAGAGCACCAGAGGAGCCCCCTTCCTGCCCAT-- 1198

Qy  1358 GCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCGTGCCCGAGAAGGAGA 1417
         |||||||||||||                          ||||||||||||||||||||
Db  1199 ----CGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCATGCTGCCCGAGAAGGACA 1254

Qy  1418 GCTGACCGTGAACGACGATCCAGAGACTGGTGGGCAAGCTGGCCAGCCAGATCT 1477
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255 GCTGACCGTGAACGACGATCCAGAGACTGGTGGGCAAGCTGGCCAGCCAGATCT 1314

Qy  1478 ACCCCGGCATCAAGGTGCGCCAGCTGCTGCGCGGCCGCCAAGGCCCTGACCG 1537
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315 ACGCCGGCATCAAGGTGAAGCAGCAGTGTGCGCAGCCCTGACCG 1374

Qy  1538 ACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGCCGAGACCGGGAGATCCTGC 1597
         |  | ||||||||||||||||||||||||||||||||||||||||||||
Db  1375 AGGTGATCCCCCTGACCGAGGAGGCCGAGCTGGCCGAGACCGGGAGATCCTGA 1434

Qy  1598 GCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGCCCGAGATCCAGA 1657
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1435 AGGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGCCCGAGATCCAGA 1494

Qy  1658 AGCAGGGCCACGACCAGTGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGA 1717
         ||||||||||||||||||||||   |||||||||||||||||||||||||||
Db  1495 AGCAGGGCCAGGGCCAGTGACCTACCAGATCTACCAGGAGCCTTCAAGAACCTGAAGA 1554

Qy  1718 CCGGCAAGTACGCCCAAGATCGCACCCGCCCACCAACGAGTGAAGCAGCTGACCGAGG 1777
         ||  | |  |||||||||||||||||||||||||||||||||||||||||||||
Db  1555 CCGGCAAGTACGCCCGCATGCGCGGCGCCCATGCGCGGCGCCCACCAACGAGTGAAGCAGCTGACCGAGG 1614

Qy  1778 CCGGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCC 1837
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1615 CCGGTGCAGAAGGTGAGCACCGAGAGCATCGTGATCTGGGGCATCTGGGGCAAGATCCCCAAGTTCAAGC 1674

Qy  1838 TGCCCATCCAGAAGGAGAGACTGGTGGACCGACTACTACTGCCAGGCCACCTGGA 1897
         |||||||||||||||||||  ||    | ||||||||||||||||||||
Db  1675 TGCCCATCCAGAAGGAGAGACTGGTGGATGGAGTACTGCCAGGCCACCTGGA 1734

Qy  1898 TCCCCGAGTGGGGAGTTCGTGTCAAACACCCCCCCCTGGTGAAGCTGTGTGTACCAGCTGGAGA 1957
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1735 TCCCCGAGTGGGGAGTTCGTGTCGAACACCCCCCCCCTGGTGAAGCTGTGTGGTACCAGCTGGAGA 1794

Qy  1958 AGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCA 2017
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1795 AGGAGCCCATCATCGTGGGCGGCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCA 1854

Qy  2018 AGATCGGCAAGGCCGGCTACGTGACCGACCGCGGCCGGCAGAGATCGTGAGCCTGACCGG 2077
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1855 AGCTGGGCAAGGCCGGCTACGTGACCGACCGCGGCCGGCAGAAGGTGTGTGAGCCATCCCCG 1914

Qy  2078 AGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGAGAGCGGCGCA 2137
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1915 AGGGGCTGCTGCGCCCGCCATCTGGGCCTGCAGGACCAGCGGCGCGCC 1974

Qy  2138 GCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCG 2197
         |||||||||||||||||||||||||||||||| ||||||||||||||||||
Db  1975 TGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCG 2034

Qy  2198 ACAAGAGCGAGGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG 2257
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2035 ACAAGAGCGAGGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG 2094

Qy  2258 TGTACCTGAGCTGGGTGCCCGCCACACAGGAGCCAGAGATCGGCCAACGAGGAGATCAAGGC 2317
         ||||||||||||||||||||||||||||||||  |  ||||||||||||||||||
Db  2095 TGTACCTGGCCTGGGTGCCCGCCCACACAGGGCATCGGCCAACGAGCAGGTGGACAAGC 2154

Qy  2318 TGGTGAGCGCCGGCATCGTCCAAGGTGCTGTTCCTGGACGGCATCGATGCGGCGCATCGTGA 2377
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2155 TGGTGAGCGCCGGCATCGTCCAAGGTGCTGTTCCTGAACGGCATCGATGCGGCGCATCGTGA 2214

Db  2378 TCTACCAGTACATGGACGACCTGTACGTGGGCGAGCGGCGGCCCTAGGATCGATTAAAAGC 2437
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2215 TCTACCAGTACATGGACGACCTGTACGTGGGCGAGCGGCGGCCCTAGGATCGATTAAAAGC 2274

Qy  2438 TTCCCGGGCTAGCACCGGTGAATTC 2463
         |||||||| ||||||||||||||||
Db  2275 TTCCCGGGGCTAGCACCGGTGAATTC 2300
```

RESULT 4
US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 4319
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6

Query Match   78.9%;   Score 1942.6;   DB 4;   Length 4319;
Best Local Similarity   89.7%;   Pred. NO. 0;
Matches 2129;   Conservative   0;   Mismatches   225;   Indels   19;   Gaps   4;

```
Qy    33 GGCCACCAGCGCCAACATCCTGATGCAGCGGCAGCAACTTCAAGGGCCCAGCGCATCAT   92
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1122 GACGAACCCCGGGACCCATCATGATGCAGCGGCGGCAACTTCCGCAACCAGCGGCGAAGCACCGT 1181

Qy    93 CAAGTGCTTCAACTGCGGCAAGGAGGGGCACCATCGCGGCAACTGCCGGCCCCCGCAA  152
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1182 CAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCGCGGCAACTGCCGGCCCCCGCAA 1241

Qy   153 GAAGGGCTGCTGAAGTGGGCAAGGAGGGGCAAGGACCACGGAGGACTGCACCGAGCGCCA  212
         || ||||||||||||||||||||||||     |||||||||||||
Db  1242 GAAGGGCTGCTGCCGCTCCGGCGCCGAAGGACACCAAATGAAAGATTGCACTGAGACACA 1301

Qy   213 GGCCAACTTCTTCCGCAGAGGACCTGGCCTTCCCCCAGGCCGAGGTCCGAGTTCCCCAG  272
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1302 GGCTAA-TTTTTTAGGGAAGATCGGCCCTTCCTACACAGGGGAAGGGCCAGGGAATTTTCTTC 1360

Qy   273 CGAGCAGAACCGCGCCAACGGCCCACCAGCCGGAGCTGCAGGTGCGCGGCG-----A  326
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1361 AGAGCAGACCAGAGCCAACAGCCCACCAGCCGAGCTGCAGGGCA------CCCTGAACTTCCCCCA  1420

Qy   327 CAACCCCGAGCGAGGCCGGGCCGAGCCGCCAGGGCA------CCCTGAACTTCCCCCA  380
         ||||||||||||||||||||||||||||||||||
Db  1421 CAACTCCCTCTCAGAAGCAGGAGCCGATAGACAAGGAACTGTATCCTTTAACTTCCCTCA 1480

Qy   381 GATCACCCTGTGGCAGCGCGCCCCTGGTGAGCATCAAGTGGGGCCAGATCAAGGAAGGC  440
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1481 GATCACTCTTTGGCAACGACCCCTCGTCACAGTAAGGATCGCGGCGCCAGCTCAAGGAGGC 1540
```

```
Qy   441  CCTGCTGGACACCGGCGGGCGGCCGGACGACACCGTGCTGCTGGAGGAGATGAGCCTGCCGGCGGCAAGTG    500
          |||||||  |||| |  ||||||  ||||  ||||||||||||||||||||||||||  ||||||||||||
Db  1541  GCTGCTGGACACCGGCCGGCCGGACGACACCGTGCTGGGAGGAGATGAACCTGCCCGGCGGCAAGTG   1600

Qy   501  GAAGCCCAAGATCGGCGGGATCGGCGGCCATCGGCACCGTGCTGCTTCATCGGCACCGACCAGAT    560
          ||||||  ||||||||||  |||||||||||||  |||||||||||||  |||  ||||  ||||||
Db  1601  GAAGCCCAAGATCGGCGGGATCGGCGGCTTCATCAAGGTGCGGCCAGTGCGGCCAGTGACCGACCAGAT   1660

Qy   561  CCTGATCCGAGATCTGCGCGAGAAGGCCATCGGCCACCGTGCTGATCGGCCCCACCCCCGT    620
          |||||  ||||||  ||  ||||||||||||  |||||||  |||||||  |||||||  ||
Db  1661  CCCGTTGGTGCTGGTGGTCGGCACCGCCACCAAGGCCATCCGCTGCTGGTGGTGGCCCCACCCCCGT   1720

Qy   621  GAACATCATCGGCCGCCAACATGCTGCTGACCAGCTGGCTGCACCCTGAAACTTCCCCATCAG    680
          |||||||||||||||||  |||||||||||||||||||  ||||||||||||||||||||||
Db  1721  GAACATCATCGGCCGCCAACCTGCTGCTGACCCAGATGGCTGCACCCTGAAACTTCCCCATCAG   1780

Qy   681  CCCCATCGAGACCCTGCCCGTGAAGCTGAAGCCCGGCATGGACGCCCCCAAGGTCAAGCA    740
          ||||||  ||||||||  ||||||  |||||||  ||  ||||||||||||||  |||||||
Db  1781  CCCCATCGAGACCACGGTGCCCGTGAAGCTGAAGCCCGGGATGGACGCCCCCAAGGTCAAGCA   1840

Qy   741  GTGGCCCCTGACCGGCAAGAAGATCAACGGCCCCGCTGACCGGCCATCTGCGAGGAGATGAGAA    800
          |||||||||||||||||||||||||||||||  |||  |||||||||||||||||||||||||
Db  1841  GTGGCCCCTGACCGGCAAGAAGATCAACGGCCCCTGGTGTGGAGATCTGCGAGGAGATGAGAA   1900

Qy   801  GGAGGGCAAGATCACCAAGATCGGCCCCGAGACCCCGTGTTCGGCCAT    860
          ||||||  ||||||||||||||||  |||||||||  ||||||||||||
Db  1901  GGAGGGCAAGATCAGCAAGATCGGCCCCGAGACCCCTGTGTTCGGCCAT   1960

Qy   861  CAAGAAGAAGGACAGCACCAAGTGCGCAGCTGGTGCGCAAGCTGGTGGACTTCCGAGCTGAACAAGCG    920
          ||||||||||||||||||||||||||||||||  ||||||||||||||||  |||||||||||||||
Db  1961  CAAGAAGAAGGACAGCACCAAGTGCGCAGCTGGTGCTGCGCAAGCTGGTGGACTTCCGAGCTGAACAAGCG   2020

Qy   921  CACCCAGGACTTCTCGGAGGTGCAGCTGCAGCTGGCCATCCCCCACCCCCGGCCGGCTGAAGAAGAA    980
          |||||||||||||||||||||  ||||||||  ||||||  ||||||||||||||  |||||||||||
Db  2021  CACCCAGGACTTCTCGGGAGGTGCAGCTGCAGCTGGCCATCCCCCACCCCCGGCCGGCTGAAGAAGAA   2080

Qy   981  GAAGAGCGTGACCGTGTGGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACCAGGA   1040
          ||||||||||||||||||||||||||||||||||||||||||||||  |||||  |||||||
Db  2081  GAAGAGCGTGACCGTGTGGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACCAAGGA   2140

Qy  1041  CTTCCGGCAAGTGCTGCCGCCAGCATCCCCAGCATCAACAACGAGACCCCCGGCATCCG   1100
          ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db  2141  CTTCCGGCAAGTGCTGCCGCCAGCATCCCCAGCATCAACAACGAGACCCCCGGCATCCG   2200

Qy  1101  CTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCCAGAGCAG   1160
          |||||||||||||||||||||||  |||||||||||||||||||||||||  |||||||||
Db  2201  CTACCAGTACAACGTGCTGCCCCAGGGGCTGAAGGGCAGCCCCGCCATCTTCCAGAGCAG   2260

Qy  1161  CATGACCAAGATCCTGGAGCCTTCGCGCCCCGCCAACCCGAGATCGTGATCTACCA---   1217
          |||||||||||||||||||||||  |||||||  |||||||  ||||||||||||
Db  2261  CATGAACCAAGATCCTGGAGCCTTCGCGGCCCTCGCAAGCAGCCCGACATCGTGATCTACCAGTA   2320

Qy  1218  ---GGCCCCGTACGTGGGCACGGCAGCAGCGACCTGGCCCCCCAAGATCGGCCCCGAAGGTCAAGCTGCT   1274
             |||||||||||||||||||||||  |||||||||||||||||||||||||||||  |||||||  |||
Db  2321  CGGCCCCGTACGTGGGCACGGCAGCAGCGACCTGGCCCCCCAAGATCGGCCACACCGGCAAGAGCGACCAGA   2380

Qy  1275  GGAGCTGCGCAAGCACCTGCTCGCGTCGCTGCGCGCTGCTGCTGCTACGGACCGTGCCAGCC   1334
          |||||||||||||||||||||||||||||||||  ||||||||  ||||||||||||  ||||
Db  2441  GGAGCTGCGCAAGCACCTTCTGTGTGATGGCGTCTACGAGCTGCACCAAGTGCACCCGGACCGTGGCAGCC   2500

Qy  1335  GGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGAACCACACCGTCGACTGGACCGTGCCAGCC   1394
          |||||||||||||||  |||||||||||||||||  |||||  |||||||||||||||||||  ||||||
Db  2441  GGAGCCCCCTTCCTGTGTGATGGGCTACGAGCTGCACCCGAACCACACCGTCGACTGGACCGTGCCAGCC   2500

Qy  1395  CATCGAGCTGCCCGAGGAGGAGAGCCTGACGCTGACCGTGAACGACATCCCGAGACCTCAGAAGCTGGCAA   1454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2501  CATCATGCTGCCCGAGGAGGAGAGCGTGACGCTGACCGTGAACGACATCCCGAGACCTGAACGAAGCTGGGCAA   2560

Qy  1455  GCTGAACTGGGCGGCCAGATCTACCCCGGCATCAAGGTGAGCAGCTGTGCAAGCTGCT   1514
          |||||||||||||||||||||||||||||||||||||||||||||  |||||||||||
Db  2561  GCTGAACTGGGCGGCCAGATCTACGCCGGCATCAAGGTGAGCAGCTGTGCAAGCTGCT   2620
```

```
Qy  1515  GGCGGCCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGCCGACCTGGAGCT   1574
          ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db  2621  GCGGCGGCCACCAAGGCCCTGACCGAGGTGATCCCCCTGACCGACCGGTGATCGGACCTGGAGCT   2680

Qy  1575  GGCCGAGAACCGGCGAGATCCTGCGCGAGCCCGTGTACTACGACCCCAGCAA   1634
          ||||||||||  ||||||||||||||||||||||  |||||||||||||||||
Db  2681  GGCCGAGAACCGGCGAGATCCTGAAGGAGCCCGTGTACTACGACCCCAGCAA   2740

Qy  1635  GGAACTGGTTGGCCGAGATCCAGAGACAGGGCAGCGAGGCCAAGACCAGTGACCTACCAGATCTACCA   1694
          ||||||  |||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db  2741  GGAACTGGTGGCCGAGATCCAGAGACAGGGCAGCGAGGCCAAGACCAGTGACCTGACCTACCAGATCTACCA   2800

Qy  1695  GGAGCCCTTCAAGAACCTGAAGACTACGCCAAGACGCACCGCCCACCACAA   1754
          |||||||||||||||||||||||||||||||||||||||  |||||||||||
Db  2801  GGAGCCCTTCAAGAACCTGAAGACTACGCCCATGCGCGGCGCCCACCACCAA   2860

Qy  1755  CGAGTGAAGCAGCTGACCGAGGGCCCTGCAGAAGATCGCCCATCGTGATCTG   1814
          ||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db  2861  CGAGTGAAGCAGCTGACCGAGGGCCGTGCAGAAGATCGCGCACCGAGAGCATCGTGATCTG   2920

Qy  1815  GGGCAAGACCCCCAAGTTCCGCCTGCCCATCCCGAGTGGGAGTTCGTGAAACACCCCCTGT   1874
          |||||||||||||||||||||||||||||||||||||||||||  ||||||||||||  |||
Db  2921  GGGCAAGATCCCCAAGTTCAAGTGCCATCTCCCGAGTGGGAGTTCGTGAACAACCCCCTGGT   2980

Qy  1875  CGACTACTGGCAGGCACCTGGATCCCCGAGTGGAGTTCGTGAACAACCCCCTGT   1934
          |||||||||||||||||||||||||||||||||||||||||  |||||||||||
Db  2981  GGAGTACTGGCAGGCACCTGGATCCCCGAGTGGAGTTCGTGAACAACCCCCTGGT   3040

Qy  1935  GAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGA   1994
          |||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db  3041  GAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCGTCGTGCGGCCCGAGACCTTCTACGTGGA   3100

Qy  1995  CGGGCGGCCAACCGGGAGACCGAGATCGGCCAAGATCGGCGAAGGCCAAGACCGGGGGCCG   2054
          ||||||||||  ||||||||||||||||  ||||||||||||||||||||||||||||||||
Db  3101  CGGCGCGGCCAACCGGGAGACCAAGGTGGGCAAGGCCGGCTACGTGACCGACCGGGGCCG   3160

Qy  2055  GCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCA   2114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3161  CCAGAAGGTGGTGAGCATCGCCGACACCAACCAGAAGACCGAGCTGCAGGCCATCCA   3220

Qy  2115  GCTTGGCCCTGCAGGACAGCGGCGGCGAGGCGGAGGTGAACATCGTGACGCAGCCAGTACCCCCT   2174
          ||||||||||||||||  ||||||||||||||||  |||  |||||||||||||||  |||||||||
Db  3221  CCTGGCCCTGCAGGACAGCGGCGGCGAGGCGGAGGTGAACATCGTGACGCAGCCAGTACCCCCT   3280

Qy  2175  GGGGCATCATCCAGGCTGAAGGGCTGAGAAGGGCAGCCCCACAAGGGCATCGGA   2234
          ||||||||||||||||||||||||  |||||||  ||||||||||||||||||||
Db  3281  GGGGCATCATCCAGGCTGAGAAGGGCTGAAGGGCAGCCCCACAAGGGCATCATCGA   3340

Qy  2235  GCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGG   2294
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3341  GCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCTGGGTGCCCGGCCCACAAGGGCATCGG   3400

Qy  2295  CGGCAACGAGCAGCTGCGGTGGCAAGCTGGTGCTGTTCCTGGAA   2354
          ||||||||||||||||||||||||||||||||||||||||||||
Db  3401  CGGCAACGAGCAGCAGTGGCGGCAGCTGGTGCTGTTCCTGGAA   3460

Qy  2355  CGGCATCGATGGCGGGCCGCCATCGTGATCTACCAGTA   2387

Db  3461  CGGCATCGACGAAGGCCCAGGCCACCAGGAGAGGAGCACCAGAGAA   3493
```

```
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 80
; LENGTH: 2305
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(+).proinact.RTopt.YM

US-09-475-515-80

Query Match        76.3%;  Score 1878.8;  DB 4;  Length 2305;
Best Local Similarity  89.4%;  Pred. No. 2.5e-306;
Matches 2061;  Conservative  0;  Mismatches 232;  Indels 13;  Gaps  3;
```

This page contains a DNA sequence alignment between Query (Qy) and Database (Db) sequences, shown in paired rows with alignment bars. The numbering runs along both margins (Qy positions 1–720 / 938–1958 and Db positions 1–779 / 780–2017).

```
Qy    60 AAGATCTGGCCTTCCTACAAGGGAAGGCCAGGAAGAATTTTCTTCAGAGAGACCAGAGCCA  119
            ||||||||||||||||||||||||||||||||||||||||            ||||||||
Db   290 ACAGCCCCACCAGCCGAGCTGCAGGTGCGGCGG------ACAACCCCGCAGCGAGG       343

Db   120 ACAGCCCCACCAGCCGAGCTGCAGGTGCGGCGG------ACAACCCCGCAGCGAGG       179
                                               CCCTGAACTTCCCCAGATCACCCTGTGGCAGC
Qy   344 CCGGCGCCGAGCGCGCAGGGCA------CCCTGAACTTCCCCAGATCACCCTGTGGCAGC  397
            |||||||||||| |||||||        ||||  |||   ||||  ||||||||||
Db   180 CAGGAGCGGTGAACATCGTGACCGACGACGACGACAAGGACTGTATCTTTAACTTCCCTCAGA  239

Qy   398 GCCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCAAGGAGGGCCCTGCTGGACACCGGCG  457
            ||   ||||||||||||||||||        |||||||||
Db   240 GACCCCTGTCACAATAAGGATCGGGGGCCTGCTCGAGCACAGAGAG       299

Qy   458 CCGACGACACCGTGCTGGAGGAGATGAGCTGCCCGGCAACGAGCTGACAAGC       517
            ||||||      ||       |||||||||||||||||||||||||
Db   300 CAGATGATACAGTATTAGAAGAAATGAATTGCCAGGAAAATGGAAACCAAAAATGATAG  359

Qy   518 GCGGCATCGGCGGCTTCATCGGCACCGTGCTCGATCGGCCCCACCCCGTGAACATCGGCCGCA  577
            |||||||||||||||  |||||||||||||||||||||||||||  ||  |||
Db   360 GGGGATCGGGGGGCTTCATCAAGGTGAGGCGGATCGGCCGATCGGCGATCTCGTG  419

Qy   578 GCAAGAAGGCCATCGGCACCGTGCTCGATCGGCCCCGTGAACATCATCGGCCGCA  637
            ||| |||||| |||   ||||||||       |||||||||||||||||||||
Db   420 GACATAAAGCTATAGGTACAGTACTATTAGTAGAGACCTACACCTGTCAACATATTGGAAGAA  479

Qy   638 ACATGCTGACCCAGCTGGGCTGCACCCTGAAC TCCCATCAGCCCCATCGGACCGTGC  697
            |||||||||||   ||||||||||||||||  |||||||||||||||||||||||||
Db   480 ATCTGTTGACCCAGATCGGCTGCACCTTGAACTTCCCATCAGCCTATTGAGACGGTGC  539

Qy   698 CCGTGAAGCTGGCCATCCCACCCCGCGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG  757
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   540 CCGTGAAGTTGGAGCTGCACGGGGATTGGAAGCGGCCCCAAGGTCAAGCAGTCGCCATTGGCCCATTGACCCGAGG  599

Qy   758 AGAAGATCAAGGCCCTGACCGGCCATCTGCGAGGAGATGAGAAGAGGCCAAGATCACCA  817
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   600 AGAAGATCAAGGCCCCTTGGTGGAGATCGCACCGAGATGGAGAAGGGAGGGCAAGATCAGCA  659

Qy   818 AGATCGGCCCCGAGAACCCCTACAACACACCCGTGTTCGCCATCAAGAAGAAGACAGCA  877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660 AGATCGGCCCCGAGAACCCCTACAACACACCCGTGTTCGCCATCAAGAAGAAGGACAGCA  719

Qy   878 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG  937
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   720 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG  779

Qy   938 AGTGCAGCTGGCCATCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGGTGACCGTGC  997
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   780 AGGTGCAGCTGGCCATCCCACCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGGCGGTGACCGTGC  839

Qy   998 TGGAGTGGCGGCACGCGCTTACTTCAGCGTGCCCTGCGGACTTCGCCGAGTACACCG  1057
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   840 TGGACGTGGGCGACGCGCTTACTTCAGCGTGCCCCTGGACACTTCGCCGAAGTACACCG  899

Qy  1058 CCTTCACCATCCCAGCATCAACAACAGAGAAGCCCCGGCATCCGTACCAGTACAACGTGC  1117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   900 CCTTCACCATCCCAGCATCAACAACAGAGAAGCCCCGGCATCCGTACCAGTACAACGTGC  959

Qy  1118 TGCCCCAGGGCTGGAAGGGCGGCAGCCGCATGAGCCCTGGAGGCGCTTCAGAGACATCCTGG  1177
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   960 TGCCCCAGGGCTGGAAGGGCGGCAGCCGCCATCTTCAGAGACAGCAGCATGACCAAGATCCTGG  1019

Qy  1178 AGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCTACCAGTACGTGGGCA  1237
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1020 AGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCTACCAGTACGTGGGGCA  1079

Qy  1238 GCGACCTGAGATCGGCGACCACCGGCGCCCGCAAGATCGGCAAGCTGCGGGA  1297
            |||||||||||||||||||||||||||||||||||||||||||||||
Db  1080 GCGACCTGAGATCGGCGACCACCGGCGCCCGCAAGATCGGCAAGCTGCTGCTGC  1139

Qy  1298 GCTGGGGCTTCACCACCCCCTTCCTGTGATGG  1357
            |||||||||||||||||||||||| ||||||
Db  1140 GCTGGGGCTTCACCACCCCCTTCCTGTGCCCAT--  1197
```

```
Qy  2018 AGATCGGCAAGGCCGGCTACGTGACGACCGGCGGCCGCGAGAAGATCGTGAGCCTGACCG  2077
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1860 AGCTGGGGCAAGGCCGGCTACGTGACCGACCGGCGGCCGCGAAGGTCGTGAGCATCGCCG  1919

Qy  2078 AGACCACCAACCAGAAACCAGAGACCAGCTGCAGGGCCATCCAGCTGCCCTCAGAGACAGCAGCGGCA  2137
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1920 ACACCACCAACCAGAAGACCGAGCTGCAGGCCATCCATCCTGGCCCTGGAGGAGCAGCAGGCAGCGGCC  1979

Qy  2138 GCGAGGTGAACATCGTGACCGACGAGCGACGGCCAGTACGCCCTGGCATCATCCAGGCCGCCGG  2197
            |||||||||||||||||||||||| |||||||||||||||||||||
Db  1980 TGGAGGTGAACATCGTGACCGACCAGCAGCCAGTACCTGGCATCATCCAGGCCCGCCAGCCCG  2039

Qy  2198 ACAAGAGCGAGAGCGAGGCGAGCTGGTGAACCAGATCATCGGCCAGCTGATCGAAGAGGAGAAGG  2257
            |||||||||||||||||
Db  2040 ACAAGAGCGAGAGCGAGGCGAGCTGGTGAGCTGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGG  2099

Qy  2258 TGTACCTGAGCTGGGTGCCCGCGCCCACAAGGGCCAACAAGC  2317
            ||||||||||||||||||||||||||||||||||||||||||
Db  2100 TGTACCTGGGCTGGGTGCCCGCCCCACAAGGGCCAACGAGCAGTGGCGGCAAGAGC  2159

Qy  2318 TGGTGAGCAAGGCCATCGGCACGGCATCGGTCGTTCCTGGACGGGCATCGATGGCGGCATCGTGA  2377
            |||||||||||||||||||||||||||||||||||||||||||
Db  2160 TGGTGAGCCGCCGGCATCCCGGCAAGGTGCTGTTCCTGGACGGGCATCGATCGGTGA  2219

Qy  2378 TCTACCAGTACATGACATGACGACTGTACGTGGGCAGCGCGCGGCCCCTAGGATCGATTAAAAGC  2437
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2220 TCTACCAGTACATGACATGACGACTGTACGTGGGCAGCGCGCGCCCCTAGGATCGATTAAAAGC  2279

Qy  2438 TTCCCGGGGCTAGCACCGGTGAATTC  2463
            ||||||||||||||||||||||||||
Db  2280 TTCCCGGGGCTAGCACCGGTGAATTC  2305
```

RESULT 6
US-09-475-515-81
; Sequence 81, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 2299
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(+).proinact.RTopt.YMWM
US-09-475-515-81

```
Query Match       75.2%;  Score 1852;  DB 4;  Length 2299;
Best Local Similarity   89.0%;  Pred. No. 7.6e-302;
Matches 2052;  Mismatches 235;  Conservative 0;  Indels 19;  Gaps 4;

Qy   170 GCGGCAAGGCAGAGGGGCCACCAGATGACGACTCCACCCGAGCTGCAGGCCCAGCCAACTTCTTCCGCG  229
            |||||||||||||||||||||||||||||||||||||||||||||||||||  |||||
Db     1 GCGGCGCGGAAGAGAAGAAGAAGATTGCACTGAGAGACGACGAGAGGAGCTAATTTTT-AGGG  59

Qy   230 AGGACCTTGGCCTTCCCAGGCAGGCCCGCAGTTCCCGCGAGGGCGAGCGAGAACCCGCCA  289
            |  ||||||||      ||  |     |||
Db   1140 GCTGGGGCTTCACCACCCCCTTCCTGTGCCCAT--  1197
```

```
Qy  2438  TTCCCGGGCTAGCACCGGTGAATTC  2463
          ||||||| ||||||||||||||||||
Db  2274  TTCCCGGGGCTAGCACCGGTGAATTC  2299
```

RESULT 7
US-09-552-950-2
; Sequence 2, Application US/09552950
; Patent No. 6551248
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; TITLE OF INVENTION: Anti-Viral Vectors
; FILE REFERENCE: 674524-2004
; CURRENT APPLICATION NUMBER: US/09/552,950
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon
; OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

Query Match   67.1%; Score 1651.8; DB 4; Length 4307;
Best Local Similarity 82.5%; Pred. No. 2.7e-268;
Matches 1939; Conservative 0; Mismatches 391; Indels 19; Gaps 4;
```
Qy    33  GGCCACCAGCGCCAACATCCTGATGCAGCGCAACTTCAAGGGCCCCAAGCGCCAGCGCCATCAT    92
          |||||| | ||| ||||||| ||||||||||||||||||||  ||  ||   |||| |||| |
Db  1110  GACCAACTCCGCTACCATCATGATGCAGCGCAACTTCCGGGCAACCAACGCAAGATCGT  1169

Qy    93  CAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCGCAACTCGCCGGCAA   152
          ||||||||||||||| ||||| || |||||| | ||| ||||||||| | |||
Db  1170  CAAGTGCTTCAACTGTGGCAAAGAAGGGCACACAGCCCGCAACTGCCAGGGGCCCCTAGGAA  1229

Qy   153  GAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGGACTGAAGGACTGAGGCGGCCA   212
          ||||||| ||| ||| ||||||||| |||| |||| |||||| |||||| ||
Db  1230  AAAGGGCGCTGAAATGCGGCAAGAAGGCCACCAGGACTGTACGGAGGACA  1289

Qy   213  GGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCAGGCCGAGTCCCCAG   272
          |||||| || ||||||||| |||||||| || |||| |||||| |||
Db  1290  GGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCTTCCTACAAGGAGGAGGAATTTTCTTC  1348

Qy   273  CGAGCAGAACCCGCCAACAGTCGTGAGCCTGACCG                 -----GGCGGGA   326
          ||||||||| | |||||||||||||||||||||                        ||||
Db  1349  AGAGCAGACCCGCGGCCACCAGCGCCGGCGAGCTTCAGGTCTGGGGTCGCGACAA  1408

Qy   327  CAACCCCCGCAGCGAGCGCCCCCCGTGCGCGAGGCCGCCA------CCCTGAACTTCCCCCA   380
          ||  |||| ||||||||||| |  |||||||||||| ||       |||||||||||||
Db  1409  CAACTCCCCCTCCGAAGCAGGGCACCGGTGTCGTCCTTCAACTTCCCTCA  1468

Qy   381  GATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGCGGCGGCCAGATCAAGGAGGGC   440
          |||||||| |||||| | ||| |||||||||| ||||| ||||| ||||| |||||||||
Db  1469  GGTCACGCTTTGGCAGCGACCCCTCGTCACCATCAAGATCGGGGGACAGTCAAGGAGAGC  1528

Qy   441  CCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTG   500
          |||||| |||||| |||||||||||||||| ||||||||||| |||||| |||||||||
Db  1529  TCTCCTGGACACCGGCGCAGACGACACCGTGCTGGAGGAGATGTCGTTGCCAGGCCGCTG  1588

Qy   501  GAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGAT   560
          |||||||| ||||| ||||||||||| |||||||||||||| |||| ||||||| ||||
Db  1589  GAAGCCGAAGATGATCGGGGGAATCGGCGGTTTCATCAAGGTGCGCCAGTATGACCAGAT  1648

Qy   561  CCTGATCGAGATCTGCGGCCACAAGGAGATCGGCACCGTGCTGGTCGGGCCCACCCCGT   620
          |||||||||| ||| ||||||||||| ||| ||||||||||| ||  ||||||||||
Db  1649  CCTCATCGAAATCTGCGGCCACAAGGCTATCGGTACCGTGCTCGTCGGGCCCACACCCGT  1708

Qy   621  GAACATCATCGGCCGCAACCTGCTGACCCAGATTAGGGGATCGATTAAAAGC   680
          ||||||||||||||||||  |||| |||| |||| |||| |||||||||||
Db  1709  CAACATCATCGGACGCAACCTGTTGACGCTGTTGCACGGTTGAACTTCCCATTAG  1768

```
(alignment continued — lower half of page)
```

Qy 1358 GCTACGAGCTGCACCCGACCAAGTGGACCGTGCAGCGCGTGCCCGAGAAGGAGA   1417
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db 1198 ----CGGAGCTGCACCCGACCAAGTGGACCGTGCAGCGCGTGCCCGAGAAGGAGACA 1253

Qy 1418 GCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCT   1477
          |||||||||||||| || ||||||||||||||||||||||||||||| |||||||||||
Db 1254 GCTGGACCGTGAACATCCAGAAGAGCTGTGTGGGCAAGCTGTGCTGCTGGGCCAGCCAGATCT 1313

Qy 1478 ACCCCGGCATCAAGGTGCGCCAGTGCGCGCCAGCTGGCGTGAGCGTGTGCAAGGCCCTGACCG   1537
          ||||||||||| |||||| | ||||| |||||||| ||| |||||| |||||||||||||
Db 1314 ACGCCGGCATCAAGGTGTGCAAGGTGTGCTGCGGCCACCAAGGGCCCTGACCG 1373

Qy 1538 ACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGC   1597
          ||| | |||||||||||||||||||| |||||||||| || |||||||||||||||||||
Db 1374 AGGTGATCCCCTACCGAGGAGGCCGAGCTGGCCGAGAACCGCGAGATCCTGA 1433

Qy 1598 GCGAGGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGA   1657
          |||||||||||||||| |||||||||||||||||||||||||| ||||||||||||| |||
Db 1434 AGGAGCCCGTGCACGAGGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGA 1493

Qy 1658 AGCAGGGGCCACGACCAGTCACCTACCAGATCTACCAGAGCCCTTCAAGAACCTGAAGA   1717
          ||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||||
Db 1494 AGCAGGGGCCAGGGGCCAGTCACCTACCAGATCTACCAGAGCCCCTTCAAGAACCTGAAGA 1553

Qy 1718 CCGGCAAGTACGCCAAGATGCGCACCGCCCACCAACGCCGCCAGACGTGAAGCAGCGTGACCGAGG   1777
          ||||||||||||||||||| || |||||||| ||||||||||| ||||||||||||||||||
Db 1554 CCGGCAAGTACGCCCGCATGCGCGCGCCGCCCACGCCCACACGCGTGAAGCAGCGTGACCGAGG 1613

Qy 1778 CCGGTGCAGAAGATGCCGCCATGGCCATGTGATCTGTGGGCCAAGACCCCCAAGTTCGCC   1837
          || |||||||||||| ||||||||||||||| ||||||||||| |||||||||||| ||
Db 1614 CCGTGCAGAAGGTGAGCGGACGGCATCGGCTACCATCTGTGGGCCAAGATCCCCCAAGTTCAAGC 1673

Qy 1838 TGCCCATCCAGAAGGAGAGACCTGGGAGACCTGGCAGGCCACCTGGA   1897
          ||||||||||||| |||||||||||| |||||||||||| |||||||
Db 1674 TGCCCATCCAGAAGGAGAGACCTGGGAGACCTGGCAGGCCACCTGA 1733

Qy 1898 TCCCCCGAGTGGAGTTCGTGAACACACCCCCCCCCTGGGTGAAGCTGTGGTACCAGCTGGAGA   1957
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db 1734 TCCCCGAGTGGAGTTCGTGAACACACCCCCCCCCTGGGTGAAGCTGTGGTACCAGCTGGAGA 1793

Qy 1958 AGGAGCCCATCATCGGCGCGCCGAGACCTTCTACGTGGACGGCGCGCCAACCGCGAGACCA   2017
          |||||||||||||| ||||||||||||||| ||||| ||||||||||||||||||||||||
Db 1794 AGGAGCCCATCATCGTGGGCGCCGAGACCTTCTACGTGGACGGCGCGCCAACCGCGAGACCA 1853

Qy 2018 AGATCGGCAAGGCCGCGGCTACGGTACCGACCGGGCGCAGAGATCGTGAGCTGACCG   2077
          |||||||| ||||||||||||||||||||||||||| |||| |||||||||||||||||
Db 1854 AGCTGGGCAAGGCCGGCGCCGCGGCTACGGTACCGACCGGGCGCAGAGGTTGGTGAGCATCGCCG 1913

Qy 2078 AGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCA   2137
          |||| ||||||| ||| |||||| ||||||||||||||||| | |||| ||||| ||| |
Db 1914 ACACCACCAACCAGAAGACCGAGCTGCAGGCCATCCACCTGGCGCTGCAGGACAGCGGCC 1973

Qy 2138 GCGAGGTGAACATCGTGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCCCAGCCCG   2197
          ||||||||||||| |||||||| |||||| ||||||||| |||||| |||||||||||||
Db 1974 TGGAGGTGAACATCGTGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCCCAGCCCG 2033

Qy 2198 ACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGGAGAGAAG   2257
          ||||||| || ||||||||| ||||||||| |||||| ||||||||||||||||| | |
Db 2034 ACAAGAGCGAGAGCGAGCTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAGG 2093

Qy 2258 TGTACCTGGAGCTGGCCCCCACAAGGGGCATCCCCACAAGGGCATCGGCGCAATCGACAACGC   2317
          ||| |||||||| ||||| ||||| |||||||| ||||||||||||| |||||||||| |||
Db 2094 TGTACCTGGCCTGGGTGCCCCCACAAGGGCATCGGCGGCATCGGCGGCATGATCGGGGGACAAGC 2153

Qy 2318 TGGTGAGCAGGGCATCCGCGCAAGTGCTGTTCCTGGACGGCATCGATGCGCACCCCGTGA   2377
          | |||||||| ||||||||| ||||| ||||| ||| |||||| ||||||| ||||||||
Db 2154 TGGTGAGCAGGGCCGCCGGCGCAAGGTGTGTTCCTGAACGGGATCGATGCGGGCATCGTGA 2213

Qy 2378 TCTACCAGTACATGGACGACGAGCTGTACGTGGGCTCCGACCTGGAGATTAAAAGC   2437
          |||||||||||||||| |||||||| || |||||| ||||| ||||| |||||||
Db 2214 TCTACCAGTACATGGATGACGACGAGCTGTACGTGGGCTCGGACCTGGAGATTAAAAGC 2273

```
Qy   681 CCCCATCGAGACCTGCCCGTGAAGCTGAAGCCGGCATGGACGGCCCCAAGGTGAAGCA  740
         |||  ||||||   ||| ||||||||||||||  |||||||||||  |||||| ||||||
Db  1769 CCCTATCGAGACGGGTACCGGTGAAGCTGAAGCCGGATGGACGGCCCGAAGGTCAAGCA 1828

Qy   741 GTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGGTGAGGAGATGGAGAA  800
         ||| || || ||||| ||||||||||| ||||  || ||
Db  1829 ATGGCCATTGACAGAGGAGAAGATCAAGGCCACCGAGACCCCGTGTTGCACAGAGATGAAAA 1888

Qy   801 GGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACACACACCCGTGTTCGCCAT  860
         ||   ||||| | |  |||||   |||| | |||||| | ||| ||  ||||||||| ||||
Db  1889 GGAAGGGAAAATCTCCAAGATTGGGCCTGAGAACCCTGGACTGGTGTTCGCCATTGTTCGGT 1948

Qy   861 CAAGAAGAAGGACACCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCG  920
         ||||||||||||   |||||| ||  |||  |  | || ||| ||||| || ||| |||| |
Db  1949 CAAGAAGAAGGACACGATCGACGCGAAAATCTCCAAGATTGGTGTGACTTCCGCGAGCTGAACAAGCG 2008

Qy   921 CACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCGGCCCTGCTGCCGAGGAAGAA  980
         ||| |||||||||  |||  |||| | ||||||   |    || ||||||| ||||| |||||
Db  1949 CACGCAAGACTTCTGGGAGGTTCAGCTGGGCATCGGCGACGCGAAAATCGGTGTGACTTCCGCGAGCTGAACAAGCG 2008

Qy   981 GAAGAGCGGTGACCGTCTGGGGGACGTGGGGTGATGGTGGGGCCTCAGGGCTGGAAGAA 1040
Db  2069 GAAATCCGTGCAAGTACACCGCCTTCACCATCCCACAGACACCCCGGCATCCG 2128

Qy  1041 CTTCCGCAAGTACAACGTGCTGCCCCAGGCTGGGAGGGCAGCCCCAGCATCTTCCAGCAG 1100
Db  2129 CTTCAGGAAGTACACTGCCTTCACAGCTGCGGTGTGGGGTCTGCCCAGTGCACCCGGGATTCG 2188

Qy  1101 CTACCAGTACAACGTGCTGCCCAGGGCTGGAGAGGCGCTCTCCCGCAATCTTCCAGAGTAG 1160
Db  2189 ATATCAGTACAACGTGCTGCTGCCCAAACCCCGAGATCGTGATCGTCATCTATCAGTA 2248

Qy  1161 CATGACCAAGATCCTGGAGCCTTCGGGCAGCAGGATCGGCAGCCTGCGCGCCAAGATCGA 1217
Db  2249 CATGACCAAAATTCCGGAGCCTTTCGGCAAACAGAACCCCGGCAGCATCGTCATCTATCAGTA 2308

Qy  1218 --GGCCCCCTGTACGTGGGCAGCGATGCGGACTCGGAGATCGGAGATAGGGCACACCCCCA 1274
Db  2309 CATGGATGACTTGTACGTGGGGTCTGATCTAGAGATAGGGCACACCCCCACAGAAGAAGCACCAGAA 2368

Qy  1275 GGAGCTGCGCGAAGCACCTGCTGCTGGGCTTCACCACCCCGACAAGAAGCACCAGAA 1334
Db  2369 GGAGCCCGCGCGCCAGCAGCACCACCCCGACAAGAAGCACCAGAA 2428

Qy  1335 GGAGCCCCCCTTCTGTGGATGGGCTACGAGCTGCACAGCCCCGACAGTGACCGTGCAGCC 1394
Db  2429 GGAGGCCTCCCTTCCTCTGGATGGGTTACCGGGTTACCGACCGTGCGAGCC 2488

Qy  1395 CATCGAGCTGCCCGGAAGGAGAGAGAGCTGGAACGAGCATCCAGAAGCTGGTGGGGCAA 1454
Db  2489 TATCGGTGCTGCCAGAGAAGAAGACAGTGGACTGTGTCAACGCAACAGAAGCTGGTGGGGGGA 2548

Qy  1455 GCTGAACTGGGCCAGGCCAGCAGCATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCT 1514
Db  2549 GTTGAACTGGGCCCAGTGACTTCACCAGGGATTTAAGGTAGGCAGGCAGGCTGTGCAAACTCCT 2608

Qy  1515 GCGCGGCCGCGCCAAGGCCGCGGAGGGCCTGACGCGAGGAGGCCGAGCTGGGAGCT 1574
Db  2609 CCGCCGGAACCAGGCCACTCACAGAGGTGATGCTCCCCTAACCGAGAGGAGGCCGAGCTGGGAGCT 2668

Qy  1575 GGCCGAGAACCTGCCCGAGATCCTGCGCGGGCTCCGAGGCGCTGTCTACCGGACCCCAGCAA 1634
Db  2669 GGCAGAAAACCGAGATCCTAAAGGAGATCCTCACCTGTGCAGCGGCGGCGGTGCTGTACTACTTTACCA 2728

Qy  1635 GGACCTGGTGCCGGGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACCA 1694
Db  2729 GGACCTGATCGGCATCGGCCGAGATCCAGCAGCAGGCCCATTCAGATTACCA 2788

Qy  1695 GGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGACCGCCGCCACCACCAA 1754
Db  2789 GGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGACCGCCGCCACCACTAA 2848
```

```
Qy  1755 CGAGCTGAGCAGCTGACCGAGCGCCGTGCGAGAAGATCGCCATGAGAGCATCCTGATCTG 1814
         ||||||  |||||||||||||||||||||||||||| ||| |||||  || ||||||||||
Db  2849 CGAGCTCAGCAGCTGACCGAGCGCCGTGCGAGAGATCACCACCGAAAGCATCGTGATCTG 2908

Qy  1815 GGGCAAGAGACCCCAAGTTCCGCCTGCCATCAGAAGGAGACCTGGTGGACCTGGTGGAC 1874
Db  2909 GGGAAAGACTCCTAGTTCAAGCTGCCATCCATCCAGAAGGAAACCTGGGAAACCTGGTGGAC 2968

Qy  1875 CGACTACTGGTGCAGGCCACCTGGATCCCCGAGTGGATGGCAGTTCGTGTGACCCCCTGT 1934
Db  2969 AGAGTATTGGCAGGCCACCTGGATTCCTGAGTGGATGTGGAGTTCGTCAACACACCCCTCGT 3028

Qy  1935 GAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGA 1994
Db  3029 GAAGCTGTGGTACCAGCTGGAGAAGGAGGCCCATAGTGGGCGCCGCCTACGTGACGATACGTCACTATAACCGGTGGA 3088

Qy  1995 CGGCGCCGCCAACCGACGAGACCAAGCTGCAAGGCCGGCTACGTGACCGAGCTGCAGGCCCATCCA 2054
Db  3089 TGGGCGCCGCTAACAGGAGACTAAGCTGCAAAGCCGGATACGTGACTGAGCTGCAGGCCATTTA 3148

Qy  2055 GCAGAAGATCGTGAGCGTGACCGAGACGCGGCAGCGAGGTGAACATCGTGACCGACAGCGCCAGCTACGCGCCCT 2114
Db  3149 ACAGAAGGTTGTGTCACCCTACCACTGACACCCAGAAGCGAGGTCGAACATCGTGACAGACTCTCAGTATGCCCT 3208

Qy  2115 GCTGGCCCTGCAGGACAGCAGGCCCGACAGCGCAGGTGAACATCGTGACCGACAGATCATCGA 2174
Db  3209 CCTCGCTTTGGAGATCGGGCCAGCCCGACAGCGAGGTCGAGCTGGTCAATCAATCAGATCATCGA 3268

Qy  2175 GGGCATCATCAACCAGGCCCGACAGCGCGAGCAGAAGTGTACCTGGTGAACCAGATCATCGG 2234
Db  3269 GGGCATCATTCAAGCCCAGCCAGACCAGAGAGTGAGTTCCGAGCTGGTCAATCAGATCATCGA 3328

Qy  2235 GCAGCTGATCAAGAAGGAGAAGGTGTACCTGGATCTGTGAACCAGATCATCGG 2294
Db  3329 GCAGCTGATCAAGAAGAGGAAAAGGTCATCTCGGCCTGAGCCTGGGTACCGGCCGCCACAAAGGCATTGG 3388

Qy  2295 CGGCAACGAGCAGATGACAAGCTGTGTGAGCAAGAAGCACCAGAA 2354
Db  3389 CGGCAATGAGCAGGTCGACAAGCTGGTGTCTGGCTGCATCAGAGGAAGGTGTATTCCTGGA 3448

Qy  2355 CGGCATCGA 2363
         |||||||||
Db  3449 TGGCATCGA 3457
```

```
Query Match          66.0%;  Score 1624.6;  DB 4;  Length 9772;
Best Local Similarity 81.8%;  Pred. No. 1.1e-263;
Matches 1922;  Conservative 0;  Mismatches 408;  Indels 19;  Gaps 4;

Qy    33 GGGCACCAGCGCCAACATCCTGATGCCGCAGCGCCAGCAACTTCAAGGGCCCAGCGCCATCAT  92
Db  2217 GACCAACTCCGCTACCATCATGATGCCGCAGCGCCGGCAACTTCGGAACCACCAAGAGATGT 2276
```

```
Qy     93  CAAGTGCTTCAACTGCGCAAGGAGGCCCACATGCGCCCGCAACTGCCGCGCCCCCGCAA   152
            ||||||||||||||||||||||||||||||||||||| ||| ||| |||||||||||||
Db   2277  CAAGTGCTTCAACTGCGCAAGTGGGAAAGAGGGCACACAGCCACCGAGGCGCCCTAGGAA  2336

Qy    153  GAAGGGCTGCTGTGGAAATGTGGGAGAAGGAGGACACCAAATGAAAGATTGTACTGAGACA   212
            ||| ||| |||||||||||||||||||||||||
Db   2337  AAAGGGCTGTTGGACCCTGGCCTTCCCCAGGGGCAAGGCCCGCGAGTTCCCCAG  2396

Qy    213  GGCCAACTTCTTCCGGAGGACCTGGCCTTCCCCACACAGGGAAGGGAAGGCCAAGGGAATTTTCTTC   272
            ||||||||||||||||||||||
Db   2397  GGCTAA--TTTTTAGGGAAGGATCTGGCCTTCCTCCAAGGAAGCCAAGGAAGTTTTTCTTC  2455

Qy    273  CGAGCAGAACCGGCGCCAACAGCCCCAACAGCCCCAGAGAGAGACGTTCAGGTTTGGGGAAGAGACAA   326
            |||||||||||||||||||
Db   2456  AGAGCAGACCAGAGCCAACAGCCACCAGCCCCAACAGCCCCAGAGAGAGACGTTCAGGTTTGGGGAAGAGACAA   2515

Qy    327  CAACCCCCGCCGAGCGGCGCCGGCGCCGAGCGCAGGGCA-----CCCTGAACTTCCCCA   380
            ||||||| ||||||||||||||||||
Db   2516  CAACTTCCCTTCAGAAGCAGGAGCCCAGGATAGACAAGGAACTGTATCCTTTAGCTTCCCTCA   2575

Qy    381  GATCACCCTGTCGGCAAGAGGCCATGGCACCTGCGTGCTGATCGCCCAGATCAAGGAGGC   440
            ||||||| ||||||||||||||
Db   2576  GATCACTCTTTGGCAGGACCCCTGTCACAATAAAGATAGGGGCAGCTCAAGGAGGC   2635

Qy    441  CCTGCTGGACACCGGCCCGACGACACCGTGCTCGAGGAGGATGAGCCGCCCCCGCAAGTG   500
            ||||||||||||||||||
Db   2636  TCTCCTGGACACCGGAGCACCCGCGCTGGAGGAGATGTCGTTGCCAGGCGCTG   2695

Qy    501  GAAGCCAAGATGATCGGCCGGCCATCGGCCGCTTCATCAAGGTGCGCCAGTACGACCAGAT   560
            |||||||||||||||||||
Db   2696  GAAGCCGAAGATCGGGGAATCGGCGCGGTTTCATCAAGGTGCGCCAGTATGACCAGAT   2755

Qy    561  CCTGATCGACTTGTGGCAAGAGGCCATCGACCGGTGCTGATCGCCCACCCCCGT   620
            ||||||||||||||||||
Db   2756  CCTCATCGAAATCTCGCGGCCACAGGCTATCGGTACCGTGCTGTGGCCCCACACCCGT   2815

Qy    621  GAACATCATGGACACGGCCGCAACATGCTGACCAGTGAACTTCCCATCAG   680
            ||||||||||||||||||
Db   2816  CAACATCATCGGAACCGCAACCGTTGACGCGCAGATCGTTGACGCGCTGAACTTCCCCATTAG   2875

Qy    681  CCCCATCGAGACACCGTGCCCGTGAAGCTGAAGCCCCAAGGTGAAGCA   740
            ||||||||||||||||||
Db   2876  CCCTATCGAGCGGTGAAGCGGTACCGGTGAAGCTGAAGCGCCCGGATGAGCGGTCAAGCA   2935

Qy    741  GTGGGCCCCTGACCGAGGAGAAGAAGATCAAGGCCCCTGACCGGAGATGGAGAA   800
            ||||||||||||||||||
Db   2936  ATGGCCATTGACAGAGGACAGAAGATCAAGGCCACTGGTGGAGAGATTTGCACAGAGATGGAAAA   2995

Qy    801  GGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACACACCCCCGTGTTCGCCAT   860
            ||||||||||||||||||
Db   2996  GGAAGGGAAAATCTCCAAGATTGGCCCGGAGAACCCCTTGGAGGAGGATGGAGAA   3055

Qy    861  CAAGAGAAGGACAGCACCAAGTGGCCAAGCTCGACGAAATGGGCGAAGAGCG   920
            ||||||||||||||||||
Db   3056  CAAGAGAAGGACAGCACCAAGTGGCCAAGCTGGTGGACTTGGGACTTGGGACTTTGCGCAAT   3115

Qy    921  CACCCAGGAGACTTCTGGAGGTGCAGCGTGCAGCCATGGGCATGCGGGCATGGGCATCCCGAGAGGA   980
            |||||||||||||||||||
Db   3116  CACGCAAGAGACTTCTGGGAGGTTCAGTGGGGTCAGTGGGCATCGGAGCA   3175

Qy    981  GAAGAGGCGTGACCGTGCTGACGTGGACGTGGGTGGGTGGGTAGCGTGACTGGAGCTTCCGGTACGGGCGAGGA   1040
            |||||||||||||||||||
Db   3176  GAAATCCGGTCCCAAGTACACCACGCTCGGCGTGTGCTCCGAGACGAGACGAGAAGATGTGGAGGCAGAGGA   3235

Qy   1041  CTTCCCGGAAGTTCACACCTGGTATGGGTGTGGTCTTCACCCTATCCTTCCCGTTCCGTTCTCCTTCCCGTT   1100
            ||||||||||||||||||
Db   3236  CCTCGCTTTGCCTTGCCCTTCACAATCCTTTGATCAACAATGAGACGTCGGGGATTCG   3295

Qy   1101  CTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCTCGGAAGGGCAGCATCTTCCAGACGCAG   1160
            ||||||||||||||||||
Db   3296  ATATCAGTACAACGTGCTGCTGCTGCCCCAGGCTGGAAAGGCTCTCCCGCAATCTTCCAGAGTAG   3355

Qy   1161  CATGACCAAGATCCTGAGACCCCGCAACCCCGAGATCGATCTACCA---   1217
```

```
Db   3356  CATGACCAAAATCCTGGAGCTTGGAGCTGTGGGCAGCCAGCTGGGCAAACAGCAGACCCGACACGTCGTCATCTATCAGTA   3415

Qy   1218  ---GGCCCCCTGTACGTGGGGGCTCGCGGTGGGCTTCCGCAAACAGCAGACCCCAGCATGACGA   1274
            ||| ||| |||||| |||||||||||||||||||||
Db   3416  CATGGATGACTTGTACGTGGGCTCTGATCTAGAGATAGGGCAGCAGCCGCACCAGATCGA   3475

Qy   1275  GGAGCTGCGCAAGCACCTGCTGCTGGCGTTCACCACCCGACAAGAAGAGCACCAGAA   1334
            ||||||||||| |||||| |||||||||||||||||
Db   3476  GGAGCTGCGGCCAGCACTGTTGAGGGACTGCAGGCCAAGAGAAGAAGCACCAGAA   3535

Qy   1335  GGAGCCCCCTTCCTGTGGATGGGTACGAGCTGGACGAGCTGACCCGACCAAGTGGACCGTGCAGCC   1394
            |||||||||||||||||||
Db   3536  GGAGCCCCCTTCCTGTGGATGGGTTACGAGCTGACCCTGACAAATGACCGTGCAGCC   3595

Qy   1395  CATCGAGCTGCCGAGAGGAGGAGCTGGACCGGGACGACTGCCGTGACGACATCCAGAGCTGGGCAA   1454
            |||||||||||||||||| |||||||||||||||||
Db   3596  TATCGTGCTGCCAGAGAAAAGACAGTCTGACTGTCAACGACGTGTCACCCCTTCCAA   3655

Qy   1455  GCTGAACTGGGCGCCAGCCAGATCTACCCCGGCATCGAGGTGCGCCAGCTGTGTCGAAGTGCT   1514
            |||||||||||||||||||
Db   3656  GTTGAACTGGGCGGCCAGTGCAGTTACCCAGGGATTAGGTGAGGCAGCAGCTGTGTGCGAAACTCCT   3715

Qy   1515  GGAGCTGCGCAAGCCAAGGCCCTGACCGACGATCGTGCCCCGCCACCAGGAGGAGCTGCGGAGCT   1574
            |||||||||||||||||||
Db   3716  CCGCGGAACCAAGGCACTCACAGAGGTGATCCCCGTAACCGGAGGCCCAGCTCGAACT   3775

Qy   1575  GGCCGAGAACCGGAGATCTCGCGAGGACCGTGTACTGACACCCCAGCAA   1634
            |||||||||||||||||||
Db   3776  GGCAGAAAAACCGGAGATCCTAAAGGAGCCGTGTACTATGACCCCTCAA   3835

Qy   1635  GAACCTGGTGCGGGAGATCCAGAAGACAGGGCCACGCAGTGGACTTACCAGATCTACCA   1694
            |||||||||||||||||||
Db   3836  GGACCTGATCGCGAGGATCCAGAAGACCAGGCAGGGCAAGTGGACCATATCGAGATTTACCA   3895

Qy   1695  GGAGCCCTTCAAGAACCTGAGAACCGGCAAGATCGCACCGCCCACACCAA   1754
            ||||||||||||||||||
Db   3896  GGAGCCCTTTCAAGAAACCTGAAGACCGGCAAGATCGCGGCTGCCCACTAA   3955

Qy   1755  CGAGCTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGCCAGAGCATCGTGATCTG   1814
            |||||||||||||||||||
Db   3956  CGACGTCAAGCAGCTGACCGAGGCCGTGCAGAAGATCACCACCGAAAGCATCGTGATCTG   4015

Qy   1815  GGGCAAGAACCCCCCAAGTTCCGGCTGGATCCCCGGGTGGGGAGTTCGTGAACACCCCCCCTGGT   1874
            |||||||||||||||||||
Db   4016  GGGAAAGACTCCTAAGTTCAAGTGCCCATCCGGGTGGGGAGTTCGTGAACCTGGTGGAC   4075

Qy   1875  CGACTACTCGGCAGCCACCTGGATCCCCGGTTGTTCGCCAT   1934
            ||||||||||||||||||
Db   4076  AGAGTATTGGCAGCACTGGATTCCTGAGAGGAGGAGAGCCCATAGTGGGCCGCCCAAAACTTCTACGTGGA   4135

Qy   1935  GAAGCTGTGGTACCAGTGACTTGGAGAAGGAGCCCATCATCGGCGCGCCGAGACCTTCACGTGGA   1994
            |||||||||||||||||||
Db   4136  GAAGCTGTGTGTACCAGCTGGGACTTGGGACTTGGGACTTTCTACGTGGA   4195

Qy   1995  CGGGGCCGCCCAACCGACCTGGGCATCCCGAGCTGGACCAAGCCGGGGCCG   2054
            |||||||||||||||||||
Db   4196  TGGGGCCGGTAACAGGGGGACTCAAGGGCAAGCCGGGGCCAG   4255

Qy   2055  GCAGAAGATCGTGAGCGTTGTCTGTCTCCGACTCCACTGGACTGGAGGA   2114
            |||||||||||||||||||
Db   4256  ACAGAAGGTTGTCACCCTGACAACACCACCAACCAGAAGAGTCGTGAGCTGCAGATTA   4315

Qy   2115  GCTGGCCCTGCAGGACAGCGGCAGCGAGCATCGACACCGTACGGGCCCT   2174
            |||||||||||||||||||
Db   4316  CCTCGCTTTGCAGGCCCTTGGAGACGCTGGCTCTGGTGAACATGCCGAACATCATCGA   4375

Qy   2175  GGGGCATCATCCAGGCCGACCGAGCGGACTCTTCCAGATCATCGA   2234
            |||||||||||||||||||
Db   4376  GGGGCATCATTCAGGCCGAGCTGGTCAATCAGATCATCGA   4435

Qy   2235  GCAGTCGATCAAGAAGCAGGAGAAGGTGTACTCGGAGGTGGGTGCCCCGCCACAAGGGCATCGG   2294
```

RESULT 9
US-09-872-733A-6
; Sequence 6, Application US/09872733A
; Patent No. 6655706
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpoIBNKan containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733A-6

Query Match   63.6%;   Score 1566;   DB 4;   Length 8366;
Best Local Similarity   80.2%;   Pred. No. 6.7e-254;
Matches 1885;   Conservative   0;   Mismatches 446;   Indels   19;   Gaps   4;

```
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
;  SEQ ID NO 1
;  LENGTH: 4338
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:  Mutated Human
;  OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733A-1

Query Match             62.1%;  Score 1530.6;  DB 4;  Length 4338;
Best Local Similarity   80.1%;  Pred. No. 5.2e-248;
Matches 1905;  Conservative   0;  Mismatches 444;  Indels  28;  Gaps  8;

Qy      14 TGGCCGAGCCATGAGCCAGGCCACCAGC--GCCAACATCCTGATGCAGCGCCAACT   70
           ||||||||||||| ||||||||||||||| |||||||||||| ||||||||||||||
Db    1085 TGGCCGAGCCGATGAGCCAGGTCACCAGTGCGCCACCATAATGATGCAGAGAGCCAACT 1144

Qy      71 TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAAGGGCCACATCGCCC  130
           |||||||||||||||||||||||||||| ||||| ||||||||||||||| |||||||||
Db    1145 TCCGAACCAGCGGCGGAAGATCGTCAAGTGCTTCAATTGTGGCAAAGAAGGGCACACCGCCA 1204

Qy     131 GCAACTGCCGCGCCCCCCCCGAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGA  190
           || ||||||||||||||| ||||||||||||||||| ||||| ||||||||||| ||||||
Db    1205 GGACTGCCGCGGGGGCCCCCCGAAGAAGGGCTGTGTTGGAAATGTGGAAAAGGACACCAAA 1264

Qy     191 TGAAGGACTGCACCGAGCGCAGGCCAACTTCTTCCGCGAGGACTGGCCTTCCCCCAGG  250
           |||||||||||||||| ||| ||||||||| | ||||| ||||||||| |||||||||
Db    1265 TGAAAGATTGTACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCTACAAG 1323

Qy     251 GCAAGGCCCGCGAGTTCCCAGCGGAGCAGAACCGCGGCCACCAGCCCGCGGAGC  310
           ||||||| || ||||||||| || |||| ||||||||||||||||||||||||||
Db    1324 GGAAGGCCAGGGAATTTTCTTCAGAGCAGAGACCAGCCCCACCAGAAGAGAGAGC 1383

Qy     311 TGCAGGTGCGCG------CGACAACCCCGGACATCCCCCTCAGCACCGCAGGGCA  364
           ||||||||| ||      || ||||||||| ||| ||||| ||||||| ||||||
Db    1384 TTCAGGTCTGGGTAGAGACAACAACTCCCCCTCAGGAGCCGATGAGACAAGGAA 1443

Qy     365 ------CCCTGAACTTCCCCAGATCACCCGTGGCAGCCATCACCCTGGTGACGATCCAAGG  418
                 |||| ||| ||||| ||||||||||||| | ||| || |||||||||| |||| |
Db    1444 CTGTATCCTTTAACTTCCCTAGATCACTCTTTGGCAACGACCCCTCGTCAGTAAGGA 1503

Qy     419 TGGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGGCGAGGCCCAGGAGG  478
           || |||| ||||| ||| |||| |||| |||| |||||||||||||||||||||| ||
Db    1504 TCGGGGGGCAACTCAAGGAAGCGCTGCTGCTGGATACAGGAGCAGATGATACACAGTA 1563

Qy     479 AGATGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGGCGGGCATCGGCGGCTTCATCA  538
           | ||||| ||||| || ||||||||||| ||||| || || |||||||||||||||||||
Db    1564 AAATGAGTTGCCAGGAACGTGCTGATGATCATAGAGAAACTCGTGATCAGTAGGATCGGGC 1623

Qy     539 AGGTGCGCCAGTACGACCAGATCCTGATCGAGCTCTGCGGACAGAGGCCATCGGCGCACCG  598
           |||||||||||| ||||| |||  |||| |||| |||||| ||| |||||| |||||||||
Db    1624 AGGTGAGGCAGTACGACCACCACC---CCGTGAACATCATCGGCCGCAACATCGTGTGACCCAGATC 1683

Qy     599 TGCTGATCGGCCCCACC----CCGTCAACTAATCATCCGTGCAACATAATTGAAGAAAATC 654
           ||| ||| |||||     || |||| ||| |||||||||| ||| ||| |||| ||||||||
Db    1684 TATTAGTAGGACCTACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACCCAGATC 1743

Qy     655 GGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGGCCCTG  714
           ||||||||| |||| |||||| |||||| |||||||||||||| |||||||||||||
Db    1744 GGCTGCACCTTGAACTTCCCCATCAGCCCCATTGAGACGGTTGCCCCGTGAAGTTGAAGCCG 1803

Qy     715 GGCATGACAGGCCCCAAGGTGAAGCAGTGAGCCCTGGCCCTGACCGAGGAAGAGATCAAGGCCCTG  774
           ||||| || ||||||||||||||| ||| ||  ||| ||||| |||||||||||||||| |||||
Db    1804 GGGATGGACGGCCCCAAGGTCAGACAATGGCCATTGCAGAAGGAAGAGATCAAGGCCTTA 1863

Qy     775 ACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCCAAGATCCACCAAGATCGGCCCCGGAAC  834
           ||||| |||  ||||||||||| |||| |||||||||||||| ||||| |||||| |||||
Db    1864 GTCGAAATCTGTACAGAGATGGAGAAGGAAGGACGCAAGATCAGCAAGTCGGGCCTGAGAAC 1923

Qy     835 CCCTACACACCCCCGTGTTCCGCCATCAAGAGAAGACAGCACCAAGTGCGCAAGCTG  894
           ||||||||| |||||  |||||| ||||||||||| || |||||||  |||||||||
```

Qy   1575 GGCCGAGAACCGGCGAGATCTGCCGAGCCGTGCACGGCGGTGTACTACGACACCCCAGCAA 1634
          ||||||||||||||||||||||||||||||| ||| |||||||||||||||||||||||||
Db   3438 GGCAGAGAACCGGAGAGATCCTGAAGGAGCCAGTAGGTGTACTACGACCCCAAGCAA 3497

Qy   1635 GGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGACTTACCAGATCTACCA 1694
          ||||||| ||||||||||||| ||| ||||| |||| ||||||| || || |||||||||
Db   3498 GGACCTGATCGCGCAGAGATCCAGAAGCCAATGACCTTACCCAAATCTACCA 3557

Qy   1695 GGAGCCCTTCAAGAACCTGAAGACCCGGCAAGATGCGCCAAGATGCGCACCGCCACACCAA 1754
          |||||||||||||||| |||||||| |||||||| || || ||||||||||||||||| |||
Db   3558 GGAGCCCTTCAAGAACCTGAAGACCTGAAGACAGCAGCAGGCAAGGTGCCCACACCAA 3617

Qy   1755 CGACGTGAAGCAGCTGACCAGCCGTGCAGAAGATCGCCATGAGCCATCGTGATCTG 1814
          ||||||||||||||||||||||| ||||| ||||||| |||||||||||||||||||
Db   3618 CGATGTGAAGCAGCTGACCAGCCGTGACAGAGATCGCAGAGATCACCAGAGCATCGTGATCTG 3677

Qy   1815 GGGCAAGACCCCCAGTTCGCCGCTTCCGCCATCCAGAAGGAGACCTGGGACTGCTGGGAC 1874
          ||||||||||||||| || || || || |||||||||||||||||||||||||| |||||
Db   3678 GGGCAAGACTCCCAAGTTCAAGTGCCCATACAGAAGGAGACATGGGACTGGTTGGAC 3737

Qy   1875 CGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGGT 1934
          || ||| |||||| |||||||||||||| || ||||||||||||| |||||||||||||
Db   3738 CGAGTACTGGCAGGCCACCTGGATCCCTGAGTGGGAGTTCGTGAACACCCCCCCTGGT 3797

Qy   1935 GAAGCTGTGGTACCAGCTGGAGAAGAGGAGCCCCATCATCGCCGAGACTTCTACGTGGA 1994
          || ||||||||| ||| | ||||||||||| ||||||| ||||||||||||||||||
Db   3798 GAAACTGTGGTATCAGCTGGAGAAGGAACCATCGTGGGAGGCCGAGACCTTCTACGTGGA 3857

Qy   1995 CGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGCCG 2054
          ||||||||||||||||||||||||| ||| ||||||| ||||||||||||||| ||||||
Db   3858 TGGGGCAGCCAACAGGGAGACCAAGGTGGGCAAGGCCGGCTACGTGACCAACCGAAGGACG 3917

Qy   2055 GCAGAAGAATCGTGAGCCTGACCGAGAGCACCAACCAGAAGACTGAGCTGAGCGTGAG 2114
          ||| ||||| |||| || || || |||||||||||||||| ||||| ||| ||| |||
Db   3918 ACAGAAAAGTGGTGACCCTGGTGACACGAGCCCCACCACCACCAGAAGAGACTGAGCTGAGCCATCTA 3977

Qy   2115 GCTGGCCCCGCAGAGCACAGGCAGCGACTGGAGTGAAGTGAACATGTGACCGACCAGTACGCCCT 2174
          ||||||| ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   3978 CCTAGCTCTGCAAGAACAGCGCGACTGGAGTGAACATGTGACCAGCAGCCACCCAGTAGTACGCACT 4037

Qy   2175 CCTAGCTCTGCAAGAACAGCAGCCCAGCCCAGACAAGAGCGAGCTGGTGAACCAGATCATCGA 2234
          || |||||||||| || ||||| ||| || || |||||| |||| ||||||||||||||||||
Db   4038 GGGCATCATCCAAGCCACAGACCAATCCGAGTCGGTGTGAACCAGATCATCGA 4097

Qy   2235 GCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAGGGGCATCGG 2294
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   4098 GCAGCTGATCAAGAAGGAGAAGGTGTACCTGGGTACCCAGCACACAAAGGAATTGG 4157

Qy   2295 CGGCAACGAGCAGCAGATCGACAAGCTGGTGAGCAAGGGCCATCCGCAAGGGTGCTTCCTGGA 2354
          || ||| ||||| |||||||| ||| ||| ||| |||| ||||| |||||||||| |||||
Db   4158 AGGAAATGAACAGTAGATAAATTAGTCAGTCTGGATCCGAAGGTGCTGTTTCCTGGA 4217

Qy   2355 CGGCATCGAT 2364
          |||||||||||
Db   4218 CGGGATCGAT 4227

RESULT 10
US-09-872-733A-1
; Sequence 1. Application US/09872733A
; Patent No. 6656706
GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23

```
Db  1924 CCCTACACACTCCAGTCTTCGCAATCAAGAAGAGGACAGTACCAAGTGAGAAAGCTG 1983
Qy  1968 CATCGGCGCCGAGACCTTCTACGTGGACGCGCCGCCAACCCGCGAGACCAAGATCGCGAA 2027
              |||||| ||||||||||||||||||||| ||| |||||  ||||||| ||||||
Db  3064 CGTGGGAGCAGAGACCCTTCTACGTGGATGGGCCGCCAACAGGCAGCCAACGGCTGGCAA 3123

Qy  895  GTGGGACTTCCGCGAGCTGAACAAGCCACCAGGACTTCTCGGAGTGCAGCTGCGGCATC 954
              |||||||||||||||||||||||||  ||||||||||||||||||
Db  2028 GGCCGGCTACGTGACCACGGCCGGGCCGCGCAGAAGATCGTGAGCCTGAGCCTGACCACCAA 2087

Qy  1984 GTGGGACTTCAGAGAGCTGAACAAGAGAAGAGAGCACTTCTTGGGAAGTTCAGCTGGGCATC 2043
              |||||||||||||||||||||||||||||||||||||||
Db  3124 GGCGAGGCTACGTGACCACCGACCAGCAGACAGAAGATGTGGTGACACCACCACCAA 3183

Qy  955  CCCACCCCGCCGGCCTGAAGAGAGAAGAAGAAGCGTGACGTGCTGCGACGTGGGCCGACGCC 1014
              |||||||||||||||||||||||||||||||||||
Db  2088 CCAGAAGACCGAGCTGCAGCCATCCAGCCTGGCCGCATCATCCAGGAGCAGCCGAGGTGAA 2147

Qy  2044 CACCATCTCCGCTGGGGTTGAAGAGAAGAAGAAGTCAGTGACAGTGCTGACATGGGGTGATGCC 2103
              ||||||||||||||||||||||||||||||
Db  3184 CCAGAGAGACTGAGCTGCAAGCCATCTACCTAGCTCTGCAAGACAGCGACCTGGAAGTGAA 3243

Qy  1015 TACTTCAGGCGTGCCCCTGGACGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGC 1074
              |||||||||||||||||||||||||||||
Db  2104 CATCGTGACCGACAGCAGCCTGGGCATCATCCAGGGCATCATCCAGCCGAGCAAGAGCGA 2163

Qy  2104 TACTTCTCCGTTCCCTTGGACGGGACTTCAGGGAGTGCACTGTCCCTTCAGCTACCTAGC 2163
              ||| ||||||||||||||||||||
Db  3244 CATCGTGACAGATCCACAGTACG-CATGGGCATCATCCAAGCACAACCAGCAACCAATCCGA 3302

Qy  1075 ATCAACAACGAGACCCCCGGCTACCTCCGCTACCGTCTACCAGATCCTGGAGCCCTGCTGAAG 1134
              ||||||||||||||||||||||||||||||
Db  2164 ATCAACAACGAGACACCAGGCCATCCGCTACCAGTCTACCAGGATGAGAAGGTGTACCTGAG 2267

Qy  1135 GGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCCGC 1194
              |||||||||||||||||||||
Db  3303 GTCAGCTGTGAACCAGCCATCATCGAGCTGTGAACCAGCAGAGGAGAAAGTGTACCTGGCC 3362

Qy  2224 GGATCACCAGCAGCCATCTCTTCAAAGCAGCATGACCAAGATCCTGAGCCCTTCCGCAAGCAA 2283
              |||||||||||||||||||||||||||||||
Db  2268 CTGGGTGCCCGCCCACAAGGGCATCGGCCGACAGCAACGACAGATCCACAAGCTGGTGAGCAA 2327

Qy  1195 AACCCGAGAGATCGTGATCTACCA------GGCCCCCGTGTACGTGGGCAGGACCTGGAG 1248
              |||||||||||||||||||||||        |||||||||||||||||
Db  3363 ATGGGTACCAGCACACAACAAGGAATTGGAGAAATGAACAAGTAGATAAATTAGTCAGTGC 3422

Qy  2284 AACCCAGAGACACTCGTGATCTATCAGTACATGGACGACTGCTGCCAAGCAACAGCAGCAGCAGCAGA 2343
              |||||||||||||||||||

Qy  1249 ATC-GGCCAGCACCGCGCCGCCAAGATGCAGGAGATCGAGGAGGAGTCGCGCTCGGGGCCTT 1307
              ||| |||||||||||||||||||||||||||||
Db  2328 GGGCATCCCGAGACGTGCTGTTCCTGACGGCATCGAT 2364

Qy  2344 ATCGGGGGACGAGCAGCAGCAAGGATCGAGGAGGAGTCGCGCTGCTGTTGAGGTGGGGGACT 2403
              |||||||||||||||||
Db  3423 TCGGATCCGGAAGGTGCTGTTCCTGCTGGAGACGGGGATCCAT 3459

Qy  1308 CACCACCCCCGACAAGAAGCACCAGAAGCACCACCAGAAGCACCACCCCCTTCTTGTGGATGGGCTACGAGCT 1367

Db  2404 GACCACCAGCAGCAAGAAGCACCACCAGAAGCACCACCAGAAGCACCAGAACCT 2463

Qy  1368 GCACCCCGACAAGTGGACGACCGTGTGGCCCCATCGAGCGCTGCCGCCGAGGAGGAGGAGCTGGACCGT 1427

Db  2464 GCATCCTGACAGTGGACAGTGGACAGTGCAGCAGCTGCAGGACCAGCAGACAGCAGAGACTGGACTGT 2523

Qy  1428 GAACGACAATCCAGAAGCTCGGTGGGCAAGTGGCAACTGGACTCCACCCCGGCAT 1487

Db  2524 GAACGACAATACAGAGAGCTGTGGGCAAGTGGGCAAGTGAACTGGCGCGCCAGATCTACCCGGCAT 2583

Qy  1488 CAAGGTGCCGCCAGCTGTGTGCAAGCTGCTGCCGCGGCGCCAAGGGCCTGACCGACATCGTGCC 1547

Db  2584 CAAAGTTAGGCAGCTGTGCAAGCTGCTGCTTGAGGAACCAAGGCACTGCAGAAGTGATCCC 2643

Qy  1548 CCTGACCGAGGAGGGCGAGCTGGAGCTGGCCGAGACCTTCTGCGCGAGCCCGT 1607

Db  2644 ACTGACACAGGAAGGCAGCAGAGAGCTAGACTGGCAGAGAGAGTCCTGGAAGGAGCCAGT 2703

Qy  1608 GCACCGGCGTGTACTACGACCCCAGCAAGGAACCTGGTGGCCAGAGGACCTGGGCAGGGCCA 1667

Db  2704 CGCAGGATGAGGGGGTGCCACCACCAACCAGATGTGAGCAAGATGCTCAAACAATACTAGTGTAAACAATAAAGCAATTTTA 2763

Qy  1668 CGACCAGTGACCTACCAGATCTACCAGATCTACCAGATCTCAAGAACCGGCAAGTA 1727
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2764 AGGCCAATGGACCTACCAAATCTACCAGGAGCAGCTGAAGACAGGCCGTGCAGAA 2823

Qy  1728 CGGCCAAGAGATGCGCACCAGCTGAAGCAGTGACGTGAAGCACCGAGGCCGTGCAGAA 1787
              ||||||||||||||||||||||||||||||||||||||||||
Db  2824 CGCAGGATGAGGGGGGTGCCACCACCAACAGATGTGAGCAGTGACGTGAAGCAGTGACGTGAAAAGGCAATTTTA 2883

Qy  1788 GATCGGCATGAGAGCCATCGTGAGCTGGGCAAGACCCCCAAGTTCCGCTGCCATCCA 1847
              ||||||||||||||||||||||||||||||||
Db  2884 GATCACCACGAGACAGTCGTGGGCAAGACTCCCAAGTTCAAGGTCGCCCATACA 2943

Qy  1848 GAAGGAGACCTGGGAGACCTACTCGCGACTACTGGTGACCGACTACTGGACCGACTACTGGTG 1907
              ||||||||||||||||||||||||||||||||
Db  2944 GAAGGAGGAGGACATGGGAGGACATGGGAGGAGCACGAGTGGTGGACTGGACCGAGGAAGGA 3003

Qy  1908 GGAGTTCGTGAACACCCCCCTGGTACCAGGTGTGTGAAACTGTCGTGATCAGCTGGAAAAACTG 1967
              |||||||||||||||||||||||
Db  3004 GGAGTTCGTGAACACCCCTCCCTTGTGAAACTGTGTGAAACTGTGTATCAGCTGGAACCCCAT 3063
```

```
RESULT 11
US-09-184-418C-8
; Sequence 8, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
;   APPLICANT: Hahn, Beatrice
;   APPLICANT: Gao, Feng
;   APPLICANT: Shaw, George
;   TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
;   TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
;   FILE REFERENCE: D6287
;   CURRENT APPLICATION NUMBER: US/09/184,418C
;   CURRENT FILING DATE: 1999-11-02
;   NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 8
;   LENGTH: 9010
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
;   OTHER INFORMATION: isolate=96ZM651; 137..1621:"gag"; 1426..4425:"pol";
;   OTHER INFORMATION: 4370..4948:"vif"; 4888..5178:"vpr";
;   OTHER INFORMATION: 5159..5373-7734; 5298..5373-7734..7981:"rev";
;   OTHER INFORMATION: 5373-7734:"tat"; 7824:"vpu"; 5565..8171:"env"; 8173..8793:"nef"
;   OTHER INFORMATION: 5387..5647:"vpu"; 5565..8171:"env"; 8173..8793:"nef"
US-09-184-418C-8

Query Match             48.9%;  Score 1203.6;  DB 4;  Length 9010;
Best Local Similarity   70.0%;  Pred. No. 3.2e-193;
Matches 1650;  Mismatches 699;  Conservative 0;  Indels   7;  Gaps   2;
```

```
Qy    14 TGGCCGAGGCCATGAGCCAGGCCACCAGCCCACATCCTGATGCAGCGCAGCAACTTCA 73
             |||||||||||||||||||||||||||||||| || ||||||||||
Db  1221 TGGCTGAGCAATGAGCCAAACAAATAAGTGTAAACAATACTAGTGTCAGAAAAGCAATTTTA 1280

Qy    74 AGGGCCCCAAGCCACCATCATCAAGTTCCGCTGCCATCCA 133
             |||||||| ||||||||||||||||||| |||||
Db  1281 AAGGGAAATAAAGAGAATAGGGCCACCATATGCCAGAA 1340

Qy   134 ACTGCCCGGCCCCCCCCGCAAGAGGCCGAAGACCGGCCAAGGAGGGGCCACCAGATGA 193
             |||||||| ||||||||||||||||||||||||
Db  1341 ATTGCAGGGCCCCCTAGGAAAAAAGGGCTGTGTGGAAAATGTGGAAAATGAAATGA 1400

Qy   194 AGGAGCTGCACCGAGCGCCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 253
             |||||||||||||||||||||||||||||||||||||||||
```

```
Db   1328  ACCAGAAGGAGCCCCCTTCGTGGATGGGCTACGAGCTGACCCGACACCAAGTGGACCG  1387

Qy   2540  ATCAGAAAGACCCCCATTCTTTGGATGGGTATGAACTCCATCCTGACAAATGGACAG  2599
Db   1388  TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGAGTTGAACGATGCAACATCCAGAAGC  1447
Qy   2600  TACAGCCTATACAGCTGCGAGAAAAAGATAGTTGGACTGTTAATGATATACAGAAG  2659
Qy   1448  TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGTGCGCCAGCTGTCA  1507
Db   2660  TGGGAAAATTAAACTGGGCAAGTCAGATTCAGATTTACCGCAGGGATTAAAGTAAGG  2719
Qy   1508  AGTGCTGTGCCGCGGCGCAACCGTGTGCCCCTGACCGACATCGTGCCCCTGACCGAGG  1567
Db   2720  AACTCCTTAGGGGAGCCAAAGCACTAACAGACATAGTACCACTAACTGAAGAAGCAGAAT  2779
Qy   1568  TGGAGCTGGCCGAGAACCGGCGAGATCCTGCCGGAGACCCGTGACGGCGTGTACTACGACC  1627
Db   2780  TAGAATTGGCAGAGAACAAGGAAATTTTAAAAGAACCAGTACGTACATGGGGTATATTATGACC  2839
Qy   1628  CCAGCAAGAACCTGGTGGCCGAGATCCAGAAGCAGCAGGGCCACGACACCAGTGGACCTACCAGA  1687
Db   2840  CATCAAAAGACTTGATAGCTGAAATACAGAAACAAGGGCATGACCAATGGACATATCAAA  2899
Qy   1688  TCTACCGAGGACCCTTCAAGAACCTGAAGACCGCAAGTACGGCAAGTGCGCACCCGCCA  1747
Db   2900  TTTACCAGGAACCATTCAAAAATCTGAAAACAGTAACATCAGTGCACTAAAATTTTC  2959
Qy   1748  ACACCAACACGTGAAGCAGCTGACCGAGAAGATCGCATGAAGATGAGAGAGCATCG  1807
Db   2960  ACACTAATGATGTAAAACAGTTAACAGAGCAGTGCAAAAAATAGCCCTGGAGAGCATAG  3019
Qy   1808  TGATCTGGGCAAGACCCCCAAGTCCGCCTGCCTGCTCGGCTGCCTGGGAGACCT  1867
Db   3020  TAAATATGGGAAAGATTCCTAATTTAGACTACCCATCCAAAAGAAACATGGGAAACAT  3079
Qy   1868  GGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCC  1927
Db   3080  GGTGGACAGCTATTGGCAGCCACCTGGATCCCGAGTGGGAGTTTGTTAATACCCCTC  3139
Qy   1928  CCCTGGTGAAGCTGTGGTACCAGCTGGAAGGAGCCCATCATCGGCCCGAGACCTTCT  1987
Db   3140  TCTTAGTAAAATTATGGTACCAGCTGGAGAAGAACCCATAGTAGGAGCAGAAACCTTCT  3199
Qy   1988  ACGTGACGGCGCGCCGCCACCGGCAAGAATCGGCAAGGGCCGGCTACGTGACCGACC  2047
Db   3200  ATGTAGATGGAGCAGCAGCCAATAGGGAAACTAAATTAGGAAAAGCAGGGTATATTACTGACA  3259
Qy   2048  GGGGCCGGCAGAAGATCGTGAGCCTGACCGAGCTGACCTTCAGCCTGGGGCTGCAGG  2107
Db   3260  GAGGAAGGCAAAAAATTGTTACTCTAACTGAAACAAAATCAGAAGACTGAATTACAAG  3319
Qy   2108  CCATCCAGTCTGGCCCTGCAGCAGCGCAGCGGAGGTGAACATCGTGACCAGCAGCCAGT  2167
Db   3320  CAATTTACCTTAGCTTTGCAAGATTCAGGATCAGAATTCAGAAGTAGTAGAACTAACTGACTCACAGT  3379
Qy   2168  ACGCCCTGGGCATCATCGAGCCCCAGAGGCTGTGAAGGGCGAGAGCGAGGCGAGCTGTGAACCAGA  2227
Db   3380  ATGCCTTAGGAAATCATTCAGCACCATCCAGATAGAGTGAATCAGAGTAGTCAACCAAA  3439
Qy   2228  TCATCGAGCAGCTGATCAAGGAGCAGGAAGAGGTGTACCTGGGCTGCCCGCCCACAGG  2287
Db   3440  TAATAGAACAATTAATAAAGAACGAAACGGGTCTACCTGTCATGGGTACCAGCACATAAAG  3499
Qy   2288  GCATCGGCCGCAACCGAGCAGATCGCACAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGT  2347
Db   3500  GAATTGGAGGTAATGAACAGGTAGATAAATTAGTAAGCAAGAGGGAATCAGAGGAAAGTGCTGT  3559
Qy   2348  TCCTGGACCGGCCATCGA  2363
Db   3560  TTCTAGATGGAATAGA  3575
```

```
Db   1401  AAGACTGTACTGAGAGGCAGGCAGCTAA-TTTTTTAGGGAAAAATTTGGCCTTCCCACAAGGGA  1459

Qy    254  AGGCCCGCGAGTTCCCCAGCCAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGAGTGC  313
Db   1460  AGGCCAGGGAATTCCTTCAGACACAGCCAACAGCCCCACCAGCCCACCAGCGAGAGCTTC  1519
Qy    314  AGGTGCGCGGCGACAACCCCGCAGCGAGGCCGGGCCAGCGGCGCCAGGGCACCCTGAACT  373
Db   1520  AGGTTCGAGAGACAACCCCGCTCCGAAGCAGGCAGCAGGAGACAGGGAAGCCTTAACT  1579
Qy    374  TCCCCCAGATCACCCTGTGGCAGGCGCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA  433
Db   1580  TCCCTCAAATCACTCTTGGCAGCGACCCCTGTGTCTCAATAAAGGTAGGGGGCCAAATAA  1639
Qy    434  AGGAGGCCCTGCTGGACACCGGCGCGACCGCGCACCGTGCTGGAGGAGATGAGCCTGCCCG  493
Db   1640  AGGAGAGCCTCTCTTAGACACCGGAGCAGGTGTATACAGTACAGTGAAGAAATAAATTTGCCAG  1699
Qy    494  GCAAGTGGAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCAAGTGCGCCAGTACG  553
Db   1700  GCAAATGGAAACCAAAAAATGATAGGAGGAAATTGGAGCGGCTTTATCGAAGTAAGACAATATG  1759
Qy    554  ACCAGATCCTGATCGAGATCTGCGCGCAAGGAAGGCCATCGGCCACGCTGCTGATCGGCCCCA  613
Db   1760  ATCAAATAACCTATGGAAATTTGTGGAAAAAGGCTATAGGTACAGTATTAGTTAGGACCTA  1819
Qy    614  CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGCTGCACCCTGAACTTCC  673
Db   1820  CACCTGTCAACATAATTGGAAGAAACATATGTGACTCAGCTTGGATGCACTAAAATTTTC  1879
Qy    674  CCATCAGCCCCATCGAGAGCTGCGTGCCGTGAAGCTGAAGCCGGCCATGACGACGGCCCCAAGG  733
Db   1880  CAATTAGTCCTATTGAAAGCTGTACCAGTAGAAAATTAAAGCCAGGAATGGATGGCCCAAAGG  1939
Qy    734  TGAAGCAGTGGCCCCCTGACCAGGAAGAGATCAAGGCCCGTGACCGGCCCATCTGCGAGGAGA  793
Db   1940  TTAAACAATGGCCATTGACACAGAAGAAGAAATAAAAGCTTTAACAGCAATTTGTGAAGAAA  1999
Qy    794  TGGGAAGAGGGGCAAGATCACCCAGATCAACCAAGATCGGCCCCGAGAACCCCTACACACACCCCGTGT  853
Db   2000  TGGGAGAAGAAGGAAAAATTACAAAATTGGCCTGAAAAATCCATATAACACTCCAGTAT  2059
Qy    854  TCGCCCATCAAGAGAGGACCAGCACCGTGCTGGCACGTGGGCCAAGTGGTCAAGTTCCGGAGCTGA  913
Db   2060  TTGCCATAAAAAGAAGAAGACAGTACTACTAAGTGGCAAATTAGTAGATTTCAGGGAACTCA  2119
Qy    914  ACAAGCGCACCCAGGACTTCTCGGGAGGTGCAGTTGGCCATCCCCCACCCCGGCCGGGGCTGA  973
Db   2120  ATAAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAATACCACACCCCAGCAGGGTTAA  2179
Qy    974  AGAAGAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG  1033
Db   2180  AAAAGAAAAAAATCAGTGACAGTACTGGATGTGGGGGATGCATATTTTTCAGTTCCTTTAG  2239
Qy   1034  ACGAGGACTTCCGGCAAGTACACCCGCCTTCACCCATCCCCAGCATCAACAACGAGACCCCCG  1093
Db   2240  ATGAAAGCTTCAGGAAATATACTGCATTTACCATACCTAGTACAAACAATGAAACACCAG  2299
Qy   1094  GCATCCGCTACCAGTACAACGTGCTGCCCCCAGGTGGAAGGGCGAGGCCCCAGCATCTTCC  1153
Db   2300  GGATTAGATATACAATATAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCC  2359
Qy   1154  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATCGTGATCT  1213
Db   2360  AGAGTAGCATGACAAAAATCTTAGAGCCCTTTCAGGGCACAAAATCCAGACACTAGTCATCT  2419
Qy   1214  ACCA------GGCCCCCCTGTACGTGGGCACGGACCTGGAGATCGGCGCCAGCACGCGCCCA  1267
Db   2420  ATCAAATATATGGGATGGATCTGATGTGTAGGATCTAGAATTAGAAATAGGGCAACATAGAGCAA  2479
Qy   1268  AGATCGAGAGACTGCGCCAAGCACCTGCTGCTCGGCCTTCACCCCGACCAAGAAGC  1327
Db   2480  AAATAGAAGAGTTAAGAGAACATCTATTAAAGTGGGATTTACCACACCAGAACAAGAAAC  2539
```

```
RESULT 12
US-09-184-418C-9
; Sequence 9, Application US/09184418C
; Patent No. 6492110
GENERAL INFORMATION:
; APPLICANT: Hahn, Beatrice
; APPLICANT: Gao, Feng
; APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
SEQ ID NO 9
; LENGTH: 8972
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; OTHER INFORMATION: isolate96ZM751.3; 137..1632:gag; 1419..4435:pol;
; OTHER INFORMATION: 4380..4958:vif; 4899..5188:vpr; 5169..7814:tat;
; OTHER INFORMATION: 5308..7938:rev; 5407..5667:vpu; 5585..8128:env;
; OTHER INFORMATION: 8130..8753:nef
US-09-184-418C-9

Query Match    47.6%;  Score 1172.6;  DB 4;  Length 8972;
Best Local Similarity   69.7%;  Pred. No. 4.9e-188;
Matches 1655;  Conservative  0;  Mismatches 694;  Indels  26;  Gaps  4;

Qy    14 TGGCCGAGCCATGAGCCAGGCCGCCACCAGCTCCCAACATCCTGATGCAGCGCGACAACTTCA 73
Db  1214 TGGCTGAAGCAATGAGCCAAGTAAACACCTGGGATTAGATATCAAACATAATGATGCAGAAAAGCAATTTA 1273

Qy    74 AGGGCCCAAGCGCCATCAAGTGCTTCAACTGCGGCAAGGAGGCCACATCGCCCCGCA 133
Db  1274 AAGCCCCTAAAGAAGATTGTAAAATGTTCAACTGTGGCAGGGAAGGGCATATAGCCAGGA 1333

Qy   134 ACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGCCACCAGATGA 193
Db  1334 ATTGCAGGGCTCCTGGAAAAAGGGCTGTTGGAAATGTGAAAGGACCACCAAATGA 1393

Qy   194 AGGACTGCACCGAGCCCAGGCCAACTTCTTCCGGCAAGGAGGCCTGCCTTCCCCCAGGGCA 253
Db  1394 AAGACTGTACTGAGAGACAGGCTAA-TTTTTTAGGGAAAATTGGCCTTCCCACAGAGGGG 1452

Qy   254 AGGCCCGGGAGTTCCCAGCGAGCAG-----------AACCGGCCACAGCC 295
Db  1453 AGCCCGGGAACTTCCTTCAGAACAGACCAGAGCCAACAGCCCCACCAGCTCCACCAGCC 1512

Qy   296 CCACCAGCCCGCGAGCTGCAGGTGCGCGGCGACAACCCCCGCAGCGGCCGGCGCCGAGC 355
Db  1513 CCACCAGCAGAGAGCTTCAGGTTCGAGGACAACCCCCTGCCCCAGGAGCAGGACAGAAA 1572

Qy   356 GCCAGGGCACCCGAACTTCCCCCAGATCACCCCTGTGGCAGCGCCCCTGGTGAGCATCA 415
Db  1573 GACAAGGAACCCTAACTGGCCCTCAAATCACTCTTTGGCAGCGACCCCTTGTCTCCAATAA 1632

Qy   416 AGGTGGGCGGCCAGATCAAGGAGCCCCTGCCTGACACCAGCACCGTGCTGG 475
Db  1633 AGGTAGGGGGGCGGCAGATCAAAGGAGTCAGTGAAATAAAGGAGAAAATGGAAACCAAGGATTATTAG 1692

Qy   476 AGGAGATGAGCCTGCCCGGAAGCCAAGATGGAAGCCAAGATGATCGGCCGCATCGGCCGCTTCA 535
Db  1693 AAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGATAGGGGAATTGGAGGTTTGAGGTTTTA 1752

Qy   536 TCAAGGTGCGCCAGTAGCCAGGACCAGATCCTGATCGAGATCCTCCGGCAAGGCCATCGGCA 595
Db  1753 TCAAAGTAAGACAGTATGATCAAATACTTATACAGAAATTTGTGGAAAAAAGGCTATAGGTA 1812

Qy   596 CCGTGCTGATCGGCCACCCCGTGAACATCGGCCGCAACATCGCCCGCAACATGCTGACCCAGCTGG 655
Db  1813 CAGTATTAGTAGGACCTACACCTGTCAACATAATTGGGAGAAATATGTTGACCCAGCTTG 1872

Qy   656 GCTGCACCTGAACTTCCCCATCAGCCCATCGACCGTGCCCGTGGAGCTGAAGCCCG 715
Db  1873 GCTGCACACTAAAATTTTCCAATTAGTCCTATTGAAACTGTACCAGTAAAATTAAAGCCAG 1932

Qy   716 GCATGACGGCCCAAGTGAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGA 775
Db  1933 GAATGGATGGCCCAAGGGTCAAACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAA 1992

Qy   776 CGGCCATCTGCGAGGAGATGAGGAAGGGCCAAGATCACCAAGATCGGCCCCGAGAACC 835
Db  1993 CAGCAATTTGTGAGAAATGGAAAAGGAAAGAAAAATTACAAAAATTGGCCCTGAGAATC 2052

Qy   836 CCTACAACACTCCGGCGTTCGCCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGG 895
Db  2053 CCTACAACACTCCAGTATTGCCATAAAAAAGAAGGACAGTACTAAGTGGAGAAAATTAG 2112

Qy   896 TGGACTTCCGGGAGCTGAACAAGCCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCC 955
Db  2113 TAGATTTCAGGGAACTCAATAAAAGAACTCAGGACTTTTGGGAAGTTCAATTAGGAATAC 2172

Qy   956 CCCACCCCGCCGCCTGAAGAGGAAGAGCGCTGACCGTGCTGGACGTGGGCGACGCGCCT 1015
Db  2173 CACACCCAGCGGGGTTAAAAAAAGAAAAGTCAGTGACAGTACTGGATGTGGGGGATGCGT 2232

Qy  1016 ACTTCAGCGTGCGCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCAACCATCCCAGCA 1075
Db  2233 ATTTTTCAGTTCCTTTAGATGAAGGCTTCAGGAAATATACTGCATTCACCATACCTAGTA 2292

Qy  1076 TCAACAACGAGACCCCCGGCATCCGCCTACCAGTACACAACGTGCTGCCCCAGGGCTGGAAGG 1135
Db  2293 TAAACAATGAAACACCTGGGATTAGATATCAAATCAATAATAATGTGCTTCCACAGGGATGGAAAG 2352

Qy  1136 GCAGCCCCAGCATCTTCCAGAGCAGCAGCATGACCAGGAACAAGATCCTGGAGCCTTCCGCGCCCGCA 1195
Db  2353 GATCACCATCAATATTCCAGTAGTAGCATGACAAAAAATCTTAGAGCCCTTTAGGACACAAA 2412

Qy  1196 ACCCCGAGATCGTGATCTACCA------GGCCCCCGTACGTGGGCAGCGACCTGGAGA 1249
Db  2413 ACCCAGAAATAGTTATCTATCAATATATGGTGATGACTTGTATGTAGGATCTGATTTAGAAA 2472

Qy  1250 TCGGCCAGCACCGGCCAAGATCGAGGAGGTGCGCAGCACCTGCTGCGCTGGGCTTCA 1309
Db  2473 TAGGGCAACACAGAGCAAAATAGAGGAGTTAAGAGAACACCTATTGAGATGGGGATTTA 2532

Qy  1310 CCACCCCGACAAGAAGGCACCAGAAGAAGGAGCCCCCTTCCTGTGGATGGGCTACCAGCTGC 1369
Db  2533 CTACACCAGACAAGAAGCATCAGAAAGAACCCCCATTTCTTTGGATGGGGTATGAACTCC 2592

Qy  1370 ACCCCGACAAGTGGACGTGGCCCATCGAGCCTGCCCGCAGCGGAGGCCGGCGCCGTGA 1429
Db  2593 ATCCTGACAAATGGACAGTACAGCCTATAGTGCTGCCAGAAAAAGACAGCTGGACTGTCA 2652

Qy  1430 ACGACATCCAGAAGCTGGTGGCAAGCTGAACTGGGCCAGCTGAACTGGCCAGATCTACCCCGGCATCA 1489
Db  2653 ATGATATACAGAAGTTAGTGGGAAAATTAAACT-GGCAAGTCAGATTTACGCAGGGATTA 2711

Qy  1490 AGGTGCGCCAGCTGTGCAAGCTGCTGCCGCGGCGCCCAAGGCCCTGACCGCGACATCGTGCCCC 1549
Db  2712 AGTAAGGCAACTGTGTAAACTCCTTAGGGGAGCCAAAAGCGCTAACAGAAGTAATCAGTATTAG 2771

Qy  1550 TGACCGAGGAGGCCGAGCTGGAGCTGGCCAAGGACCTGGTGCCGCAAGATCCGGCGGCTTCA 1609
Db  2772 TGACTGAAGAGGCAGAGTTAGAATTGGCAGAGAACAGAGAGAATTCTAAAAGAACCAGTAC 2831

Qy  1610 ACGGCGTGTACTACGACCCCGGCAAGGACCTGGTCGCGGAGATCCAGAAGGCCATCGGCA 1669
Db  2832 ATGGAGTATATTATGACCCATCAAAAGACTTAATAGCTGAAATACAGAAACAAGGGCATG 2891

Qy  1670 ACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACATGCTGACCCAGCTGG 1729
Db  2892 ACCAATGGACATATCAAGTTTACCAAGAACCATTCAAAAATCTGAAAACAGGAAGTATG 2951
```

```
Qy    14  TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCACCAGCCCAACATCCTGATGCAGCGAGCAACTTCA    73
             ||| ||||||||||||||| ||||||||||||||||| |||| |||  |||||||||||||||||||
Db  1216  TGGCTGAGGCAATGAGCCACCACTAATGATGTCAAAAACAGTTAACAGAGCGGGTGCAAAAAA       1272

Qy    74  AGGGCCCCAAGCGGCATCATCAAGGTGCTTCAACTGCGGCAAGGAGCATCGCCGCGCA            133
             |||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  1273  AAGGCCCTAAAAGAGAGATTGTTAAATGCTTCAACTGTGGCAAGGGCACATAGCCAGAA       1332

Qy   134  ACTGCCGGCCCCCCCCCGGTGAAGCTGTGGTACCAGCTGGAGGAGGAGCCCATCA            193
             |||||| |||||||||||||||||||||||||||||||||||||||||||| |||
Db  1333  ATTGCAGGGCCCTGAGAAAAGAGGCTGTTGGAAATGTGGCAAGAAGGACACCAAATGA      1392

Qy   194  AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGAAGTGGCTTCCCCAGGGCA             253
             |||||||||| ||||  |||||||||   ||||||||  |||||||||||||||||||
Db  1393  AAGACTGTACTGAGAGGCAGGCCTAA-TTTTTAGGGAAAATTTGGCCTTCCCACAAGGGG      1451

Qy   254  AGGCCCGCGAGTTCCCAGCGGAGGACGAACCGCGCAACAGCCCACAGCCCGCGGAGCTGC           313
             |||||| |||||||||  |||||||||||||||||||||||||||||||| ||||||||
Db  1452  AGGCCAGGGAATTCCTTCAAAACAGGCCAGCCAACAGAGGCCAACAGCAGCAGAGAGCTTC     1511

Qy   314  AGTGCGCGGCACAACCCCGCAGCGGAGGCCGGCCGAGCCGCAGGCCACCTGAACT                373
             |||| || ||||||||| |||||||| || ||||| |||||||||||||||||||
Db  1512  AGTTCAAGGAGCAGACACCCCGCCTCCGAAGCAGGAGTCGAAGACAGGGGAACCCTTAACT    1571

Qy   374  TCCCCCAGATCACCCGTGCGACGCGCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA            433
             |||| |||| || |||  |||||||||||||||| || |||||  ||||||||||| 
Db  1572  TCCCTCAAATCACTCTTTGCAGCGACCCCCTTGTCTCAATAAAAAGTAGGGGGCCAGATAA      1631

Qy   434  AGGAGGCCCTGCTGACACCGGCCGCCGACACCGTGCTGGAGGAGATGAGCCTGCCCG              493
             || ||  ||| || ||||  |||||||| ||||||||| |||||||||||||||||
Db  1632  AGGAAGCTCTCTTAGACACAGGAGCAGATGATACAGTATTAGAAGAAATAGCTTTGCCAG     1691

Qy   494  GCAAGTGGAAGCCCCAAGATCATCGGCGCCGATCGGACATCGGCGGCTTCATCAAGGTGCCCAGTACG     553
             |||||| |||| || |||||||  ||||  ||||||||  ||||| || |||||||||||  |||
Db  1692  GAAGATGGAAACCAAAATGATAGGGGAAATTGGAGGTTTTATCAAAGTAAGACAGTATG      1751

Qy   554  ACCAGATCGATCGGAGATTCTGCGGCCAAGAAGGCCATCGGCCACCGTCGATCGTCGGCCCCA        613
             |||  ||||||||| ||||||  |||||||||||||||| |||  ||||| |||
Db  1752  ATCAAATACTTATAGAAATTTGTGGAAAAAAGGCTATAGGTACAGTATTAGTAGGACCTA     1811

Qy   614  CCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGACGCCCGGCCAACTTCC            673
             |||  || |||||||  ||||||  ||| |||||||   |||||||| |||  |
Db  1812  CACCTGTCAACATAATTGGAAGAAATATGTTGACTCAGCTTGGATGCACTTTAAATTTTC      1871

Qy   674  CCATCAGCCCCATCGAGACGTGCCCGTGAAGCTGAAGCCCGGCATGACGGCCCCAAGG            733
             |||  | ||||| ||| |||   |||| ||||||   ||||  |||||  ||||||||
Db  1872  CAATTAGCCCCATTGAAACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAGG    1931

Qy   734  TGAAGCAGTGCCCCTGACCGGAGAAGATCAAGGCCCATCTGCGAGGAGA             793
             ||||||| |||  ||||||| |||||| || |||  |||| |||||||
Db  1932  TTAAACAGTGGCCATTGACAGCAGAGAAAATAAAAGCATTAACAGAAATTTGTAAAGAAA    1991

Qy   794  TGGGAGAGGAGGAGCCAAGATCACCAAGATCGGCCCCAGAACCCCTACAACACCCCGTGT         853
             ||||| || |||| ||||||  |||  |||   || ||||  | ||||||||| |||
Db  1992  TGGAGAAGGAAGGGAAAATTACAAAAATTGGGCCTGAAAATCCATATAACACTCCAGTAT    2051

Qy   854  TCGCCATCAAGGAGAAGGACAGCACCAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA          913
             ||| || |||  ||||| |||| ||| ||||||  | |||| ||| ||  ||||  |||
Db   854  TTGCCATAAAGAAAAAGGACAGTACTAAGTGGAGAAAATTAGTAGATTTCAGGGAGGTCAA      913

Qy   914  ACAGGCGCACCCAGGACTTCTCGGGAGGTGCAGCTGGCATCCCCCACCCCCGGCCCTGA         973
             | | || || |||||||||||  ||||  | ||| ||||||| || |||||||||||
Db  2112  ATAAAAGAACTCAAGACTTTTTGGGAAGTTCAATTAGGAATACCACACCCAGCAGGTTTAA   2171

Qy   974  AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG         1033
             |||||| | | || |||| |||  ||  ||||| |||||||| ||||||||| ||||
Db   974  AAAGAAAAAATCAGTGACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCTTTAG     2231

Qy  1034  ACGAGGACTTCCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACAAGACACCCCCCG        1093
             ||||||  |||  |||||| | |  ||||||| |||||| ||  || || | || || 
Db  2232  ATGAAGGCTTCGGGAAAATATACTCGCATTCACCATACCTAGTATAAACAATGAAACACCAG    2291
```

```
Qy  1730  CCAAGATGCCGCACCGCCCACCAACGACCTGAAGCAGCAGCTGAACCGGCCGTGCAGGAAGA     1789
             |  ||| | ||| |  ||| |||  | |  | ||| |||||   ||  | ||   |||||
Db  2952  CAAAAATGAGGACTGCCCACACTAATGATGTAAAACAGTTAACAGAGCGGGTGCAAAAAA      3011

Qy  1790  TCGCCATGGAGCAGCATCGTGATCTGGGCCAAGACCCCCAAGTTCCGCCTGCCCATCCAGA     1849
             |||| | ||| || |   ||||| ||| ||||  |  ||||||| |  |||  |||||||
Db  3012  TAGCCATGGAAAGCATAGTAGTGTGGGAAGATTCCTAAATTTAGGCTACCATTCAAA      3071

Qy  1850  AGGAGACCTGGGAGACTGGTGGACCGACTACTGCGACCTGCAGGCCACCTGAAGTGGG        1909
             ||  ||||| |||| |||| || ||||| | ||| | || ||||||| |||| ||||
Db  3072  AAGAAACATGGGAGCACATGGACAGTGGACTAGATTCGCAAGCCACCTGGATTCCTGAGTGGG   3131

Qy  1910  AGTTCGTGAACAACCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGGAGGAGCCCATCA     1969
             | ||||| | ||||    ||  ||||| |||||||  ||| ||  | |||||||||||
Db  3132  AGTTTGTTAATACCTCCCCCCCTAGTAAAATTATGTACCAGCTGGAGAAAGACCCCATAG     3191

Qy  1970  TCGGGCCGGAGACCTTCTACGTGACGGCGCCGCCCAACCCGCCGCCGAGATCGGCAAGG     2029
             || | ||  || || ||| | ||| || ||| |  ||||  || ||  |||  ||||
Db  3192  CAGGAGCAGAAAACTTACTATATGTGAGATGGGCAGCCAATAGGGAAACTAAAATAGGAAAAG   3251

Qy  2030  CCGGGTACGTGACCGACCGGGGCCGCGAGAAGTCGTGAGCCTGACCGAGACCCACCAACC     2089
             || | |||||  ||||| |  || || |||| | || | |||||| |||||||||||||
Db  3252  CAGGGTATGTTACTGACAGAGGAAGGCAAAAAATTGTTACTCTAACTGAAACAACAAATC     3311

Qy  2090  AGAAGACCGAGCTGCAGGCCCATCCAGCTGGCCCCTGCCAGGACAGCGCCAGCGAGGTGAACA      2149
             || ||||||  ||||  || |||  || |||   |||||||| || |||||||| |||||
Db  3312  AAAAGACTGAATTACAAGCAATTCAGTTAGCTTTGCAGGATTCAGGATCAGAAGTAAACA     3371

Qy  2150  TCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCACAAGAGCCGGA         2209
             || ||| || || || | |  ||||| ||||| | |||  ||| | |||||| 
Db  3372  TAGTAACAGACTCACAGTATGCATTAGGAATCATCCAAGCACAACCAGATCAAGAGTGAAT    3431

Qy  2210  GCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGAGAAGGTGTACCTGAGCT    2269
             || ||||| || ||||| || ||| |||  |||| |  ||||| |||||| ||||| |
Db  3432  CAGAATTAGTCAATCAAATAAGAAACAGTTGATCGGCGGCCAACAGCACAATAAGAGTGAAT   3491

Qy  2270  GGGTGCCGCGCCACAAGGGCATCGGCCGCGGCCAACGACATCGACAAGCTGGTGAGCAAGG    2329
             |||||| | ||||||| ||||| ||  ||| || ||| ||||| |||  | || ||||
Db  3492  GGGTACCAGCAGCACAAAGGAATTGGAGGAAATGAACAAGTAGATAAATTGGTAAGTGAGTG    3551

Qy  2330  GCATCCGCAAGGTGCTGTTCCTGACGGCATCGAT    2364
             ||||| ||  ||||| ||| | |  | | |||||
Db  3552  GAATCAGGAAAGTGCTGTTTCTAGATGGAATAGAT      3586
```

```
Qy  1094  GCATCCGCTACCAGTACAACGTGCTGCCCAGGGCTCGGAAGGGCAGCCCCAGCATCTTCC  1153
          ||||| ||||| |||||| |||||||  | ||||||||| |||||| ||||| |  ||
Db  2292  GGATTAGAGATCAATATATATGTGCTTCCACAGGGATGGAAAGGATCACCCAGCAATATTCC  2351

Qy  1154  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCCGAGATCGTGATCT  1213
          || ||| | ||||| ||||||||| || ||||||||  |||||||||| ||| ||||
Db  2352  AGAGTAGCATGACCAAAAATCTTAGAGCCCTTTAGGGCGCAGAAATCCAAAAATAGTCATCT  2411

Qy  1214  ACCA------GGCCCCCGTGACGTGGGCAGCGACCTTGAGATCGGCCAGCACCCGCCA  1267
          |||        |  ||||||  ||  |||||||| ||||   || |
Db  2412  ATCAATATATGATGAGTTGTATGTAGGGTCTGACTTAGAAATAGGGCATCATAGAGCAA  2471

Qy  1268  AGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAAGC  1327
          || ||||| |||  |||||||| |||||||  | |  ||||  ||| |||| |
Db  2472  AAATAGAGGAGTTAAGAGCACATCTATTAAAGTGGGGATTCACCACCACCAGATAAGAAAC  2531

Qy  1328  ACCAGAAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGACAAGTGACCG  1387
          ||||| || || ||||  |  |  | || ||| || ||| ||  |||||  |||||
Db  2532  ATCAGAAAGAACCCCCATTTCTTTGGATGGGGTATGAATCCCATCCTGACAAAATGGACAG  2591

Qy  1388  TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACGGTGAACGACATCCAGAGCTGG  1447
          || ||||||| ||| | || ||||| || |||  |||| ||| ||||| || ||||
Db  2592  TACAGCCTATAAAGCTGCCAGAAAAGGATAGCTGGACTGTGTCAATGATATACAGAAGTTAG  2651

Qy  1448  TGGGCAAGTGAACTGGCGCCAGCGAGATCTACCCCGGCCATCAAGGTGCGCCAGCTGTGCA  1507
          ||||||||| || |||||| |||  | |||||| || |||| || ||| ||||| |
Db  2652  TGGGAAATTAAACTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGTGAGGCAACTTTGTA  2711

Qy  1508  AGCTGCTGCGCGGCCAAGGCCCTGACCGACATCTGTGCCCCTGACCGAGGAGGCCGAGC  1567
          ||| | ||||| | || || |||  | || ||  | ||| ||  | |||||
Db  2712  AACTCCTTAGGGGGCCCAAAGCACTAACAGACATAGTACCACTAACTGAAGAAGCAGAAT  2771

Qy  1568  TGGAGCTGGCCGAGAACACCGGCGATCTCGGCGAGCCCGTGCACGGCGTGACTACGACC  1627
          ||||||| ||||||||| || |||||| ||  | ||||| || || | |||  ||||
Db  2772  TAGAATTAGCAGAGAACAGGGAAATTCTAAAAGAGCCAGTACATGGAGTGTATTATGACC  2831

Qy  1628  CCAGCAAGGACCTGGTGCCGGAGATCCAGAGCAGCGGGCCACCGACCTACCAGA  1687
          |||| ||| ||| | |||  |||||| | |||| ||| ||||  | ||||||
Db  2832  CATCAAAAGACTTAATAGCTGAATACAGAAAACAGGGGCATGGTCCATGATATTCAAA  2891

Qy  1688  TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCGCCAGCAGTGCCACCGCCC  1747
          || ||||||||| ||||| |||| || ||| | ||| ||| |||  ||| |||
Db  2892  TTTACCAGAACCATTCAAAAATCTGAAAACAGGGAAGTATGCAAAATGAGGACTGCTC  2951

Qy  1748  ACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCGACGAGATCGCCATGGAGGAGCATCG  1807
          |  |||| || ||||| || ||  | ||| ||| |||| | ||  | || | ||
Db  2952  ACACTAATGATGATGGGAAATGTAAAAACAGTTAACAGAGGCAGTGCAAAAATAGAGAAGCATAG  3011

Qy  1808  TGATCTGGGGCAAGACCCCCAAGTTCGGCTGCCTGCCCAAGTCGCCAAGTGGCGCAGCGACCT  1867
          || ||||||| || ||||| ||| | |  | |  | ||| || |  |||
Db  3012  TAAATAT--GGGAAAGACCCCAAAGTTTAAACTACCCATACAAAAGGAAACATGGGAAACGTGGGACAT  3070

Qy  1868  GGTGACGACTACTGGACGAGATCCCGAGTGGAGTTCGTGAACACCCCCTC  1927
          | ||| |||||  | | ||| ||||||||  | | || |
Db  3071  GGTGGACAGAGTATTGGCAGGCCACCTGGATTCCTGATTGGGAGTTTGTTAATACCCCTC  3130

Qy  1928  CCCTGGTGAAGCTGTGGTACCAGGTGAGGAAGAGCCCATCATCGCGCCGAGAACCTTCT  1987
          | | ||| ||||| || |||| ||  |||||||| ||  |  | |||| |||| |||
Db  3131  CCCTAGTAAAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACTTTCT  3190

Qy  1988  ACGTGGACCGGCGCCAACCCGAGAGCAAGAGTCAGTGACCGACC  2047
          |  || || ||  ||||| | || |  ||| ||||| ||||||
Db  3191  ATGTAGATGGAGCAGCTAATAGGGAAACTAAATTAGGAAAAGCAGGGTATGTTACTGACA  3250

Qy  2048  GGGCCCGGCGAACGAGATCGTGAGCCTGAGACGGTGAAGGGCGCCCAGCAGCTGCAGG  2107
          |||| || |||| |||||  |   ||||||||  | |  ||| ||  | ||
Db  3251  GAGGAAGACAGAAAATTGTTTCTTTAACTGAGACAACAAATCAGAAGACTGAATTGCAAG  3310

Qy  2108  CCATCCAGCTGGCCTGGCCTGAGGACGACGGCCAGCAGTCGAACATCGTGACCGACCAGT  2167
          | ||  |||  | || |||||  |||  |||  |  |||||||  |||| ||
Db  3311  CAATTCAGCTAGCTTTGCAAGATTCAGGAACAAGAAGTAAACATAGTAACAGACTCACAGT  3370

Qy  2168  ACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGGCGAGAGCGGAGCGACTGGTGAACCAGA  2227
          |||| ||||||| | || ||| || ||  | |  |||| | |||  ||| |||
Db  2168  TTTCAAAGGACATGGACCATGACCAAGATCCTGAGCGCCCTTTTCCAAGCAAAACCCAGATACATGGTG  2227
```

```
Db  3371  ATGCATTAGGAATCATTCAAGCACAACCAGATAAAAGTGAATCAGAGTTAGTCAACCAAA  3430

Qy  2228  TCATCGACCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGCGTGCCCGCCCACAAGG  2287
          |||| ||  |  ||| ||||| || |  | | | |||  | |||| |  | |||| |||
Db  3431  TAATAGAACAATTAATAAACAAGGAAAAGAGTCTATCTGTCATGGGTACCAGCACATAAAAG  3490

Qy  2288  GCATCGGCGGCAACGACGAGCAGATGACAAGCTGGTGACAAGCTGGTGGGCATCCGCAAGGGTGCTGT  2347
          |  || ||  | || || |||  | | | ||| |  |||| |||| | ||||  | |||  |||
Db  3491  GAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTAGTGGAATTAGGAAAGTACTGT  3550

Qy  2348  TCCTGGACGGCATCGAT  2364
          ||||| |||||  ||
Db  3551  TTCTAGATGGGATAGAT  3567
```

RESULT 14

US-09-872-733A-3

```
; Sequence 3, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
;   APPLICANT: The Government of the United States of America, as
;   TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
;   TITLE OF INVENTION:   SIV ENV GENES
;   FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
;   CURRENT APPLICATION NUMBER: US/09/872,733A
;   CURRENT FILING DATE:  2001-06-01
;   PRIOR APPLICATION NUMBER: PCT/US00/34985
;   PRIOR FILING DATE:  2000-12-22
;   PRIOR APPLICATION NUMBER: 60/173,036
;   PRIOR FILING DATE: 1999-12-23
;   NUMBER OF SEQ ID NOS: 19
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 2467
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Mutated Human
;   OTHER INFORMATION: Immunodeficiency Virus - 1 Pol gene
US-09-872-733A-3
```

```
Query Match             47.3%;  Score 1163.8;  DB 4;  Length 2467;
Best Local Similarity   84.0%;  Pred. No. 1.2e-186;
Matches 1328;  Conservative    0;  Mismatches 247;  Indels    6;  Gaps    1;

Qy   790  GAGATGGAGAAGGAGGAGGGCAAGATCACCAAGATCGGCCCGAGAACCCTACAACACCCCC  849
          ||||| ||||| || |||  ||| || |  |||||||| || || ||  || ||| |
Db     7  GAGATGGAGAAGGAGGAGGAAGATCAGGAAGATCAGGCAAGATCGGCTGAGAACCCTACAAACACTCCA  66

Qy   850  GTGTTCGCCATCAAGAAGCAGCACCAAGTGCGCAAGCTGGTGGACTTCCCGGAG  909
          |||| |||||| |||||| || |||||  |||   ||||||||  |||||||||
Db    67  GTCTTCGCAATCAAGAAGAAGGACAGTACCAAGTGCTGGTGGACTTCAGAGAG  126

Qy   910  CTGAACAAGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCCGGC  969
          |||||||||| | ||||| ||||||||| | ||||||||||  |||  | || ||  |
Db   127  CTGAACAACAGAGACTCAGGACTTCTGGGAAGTTCAGCTGCGCATCGGGCCATCCCGGCTCGGG  186

Qy   970  CTGAAGAAGAAGAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCC  1029
          |||| || ||| || |||  ||  |  ||| |||  ||| ||| | |||  ||  | |||
Db   187  TTGAAGAAGAAGAGTCAGTGACAGTGCTGCTGATGTGGGTGATGTCCTTATCTCCGTTCCC  246

Qy  1030  CTGGACGAGGACTTCCGGAAGTACACCGCCTTCACCATCCCAGCATCAACAACAGAGACC  1089
          |||||| |||||| | | || ||||||  || |||||||||||||| |||  || |||||
Db   247  TTGGACGAGGACTTCAGGAAGTACACTGCCTTCACGATCCCACGATACCTAGCATCAACAACGACA  306

Qy  1090  CCCGGCATCGACTACCAGTACACAACGTGTCGCCCAGGGCTGCGAAGGGCAGCCCCAGCATC  1149
          ||| | ||| | |||||| || |||  ||  |||  | ||  |  | || ||||||||||  |
Db   307  CCAGGCATCCGCTACCAGTACACACGTGCTGCCCACAGGGATGGGAATGGAAGGGATCACCAGCATC  366

Qy  1150  TTCCAGAGCAGCATGACCAAGATCCTGAGCGCCCTTCCGCAAGCAAAACCCAGATACGTG  1209
          ||||| ||| |||||||| |||| | |||  || |||  |  || | ||||||||| ||
Db   367  TTTCAAGGACATGGACCATGACCAAGATCCTGAGCGCCCTTTTCCAAGCAAAAACCCAGATACGTG  426
```

```
Qy   2284  AAGGGCATCGCGCGCAACGAGCGAGATCGACAAGCTCGGTGAGCAAGGGCATCCGCAAGGTG  2343
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1507  AAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGGATCCTGGGAAGGTG  1566

Qy   2344  CTGTTCCTGACGGCATCGAT  2364
           ||||||||||||||||||||
Db   1567  CTGTTCCTGACGGGATCGAT  1587

RESULT 15
US-09-184-418C-4
;  Sequence 4, Application US/09184418C
;  Patent No. 6492110
;  GENERAL INFORMATION:
;    APPLICANT:  Hahn, Beatrice
;    APPLICANT:  Gao, Feng
;    APPLICANT:  Shaw, George
;    TITLE OF INVENTION:  CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
;    TITLE OF INVENTION:  IMMUNODEFICIENCY VIRUS TYPE 1
;    FILE REFERENCE: D6287
;    CURRENT APPLICATION NUMBER: US/09/184,418C
;    CURRENT FILING DATE: 1999-11-02
;    NUMBER OF SEQ ID NOS: 112
;  SEQ ID NO 4
;    LENGTH: 8992
;    TYPE: DNA
;    ORGANISM: Human immunodeficiency virus type 1
;  FEATURE:
;    OTHER INFORMATION: isolate=92RW009; 139.1624:gag; 1690.4428:pol(N-terminus uncertail
;    OTHER INFORMATION: 4373.4951:vif; 4891.5181:vpr; 5162.7801:tat; 5301.7958:rev;
;    OTHER INFORMATION: 5403.5648:vpu; 5566.8148:env; 8150.8773:nef
US-09-184-418C-4

Query Match        46.4%;   Score 1142;   DB 4;   Length 8992;
Best Local Similarity   68.6%;   Pred. No. 6.5e-183;
Matches 1620;  Conservative  0;  Mismatches 730;  Indels  10;  Gaps  3;

Qy     14  TGGGCCGAGGCCATGAGCGCCACCAGCGCCACCATCCTGATGCGCGCAGCGCAGCAACTTCA    73
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1221  TGGGTGAAGCAATGAGCGCAAGGTACAACAACCAAACAACATATATGAGAGGCAATTTTA  1280

Qy     74  AGGGCCCAAGCGCCATCAAGTCGCTTCAAGTGCTTCAACTGCGCAAGGGCCAAGCATCGCCGCA   133
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1281  AGGGCCAGAAGAGAATTATTAAGTGTTTCAACTGTGGGCAAAGAAGGACCACCTAGCCAGAA  1340

Qy    134  ACTGCGCGCCCCCGCAAGACCGGCAAGTACCGGCAAGATGCCACC   193
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1341  ATTGCAGGGGCCCCTAGAAAAAAGGGCTGTTGGAAAATGTCGGAAGAGAGGAGGACCACCAAATGA  1400

Qy    194  AGGACTGACCAGCGAGGCCAACTTCTCGCGGAGGATCGCGAGACC   253
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1401  AAGAACTGCACTGAGAGACAGGCTAA-TTTTTTAGGGAAAATTGGCCTTCAACAAGGGGG  1459

Qy    254  AGGGCCCGCGAGTTCCCCAGCGAGCAGAGACCGGGCCAACAGCCCGCGAGCT--   311
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1460  AGGGCCAGGAAATTTTCCCCAGCGAGACTGAGCGAGGCCAACAGCCCACCGCAGAGAACTTT  1519

Qy    312  -GCAGGTGCGCGGCCAGAAGATCGTGACGACCACCCCCCGCAGCGAGCGCCAGCGGCACCCTGA   370
           |||||||||||||||||||||||||||||||||||||||||||
Db   1520  GGAATGGGGAAGGAAAGACAGATAGCCTCTTCCTCTGAAACAGGAGCAGCAGGGAACCTTTA  1579

Qy    371  ACTTCCCCCAGATCACCCTGTGCGACGACTGGAGCATCAAGTGGTGGCGGCCAGA   430
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1580  ATTTCCTCAATCATCTCTTGGCAACGACCCCCTGTCACAGTAAAAATAGGAGGTCAGC  1639

Qy    431  TCAAGGAGGCCCTGCTGGACACCGAGTGGTGGAAC   490
           |||||||||||||||||||||||||||||||||||||
Db   1640  TAAGAGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATAATGTTGC  1699

Qy    491  CCGGCAAGTCGGAAGCCCAAGATGATCGGCGGCTTCATCAAGTGTGCGCCCAGT   550
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1700  CAGGAAAATGAAACCAAAAATGATAGGGGGAATTGGGGGTTTTATCAAGGTAAAAACAGT  1759
```

```
Qy   1210  ATCTACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGC  1263
           |||||||        |||||||||||||||||||||||||||||||||||||
Db    427  ATCTATCAGTACATGGACGACATGGACCCTCACGTACGGTAGGAAGTGACCTGGGCAGCACAGG   486

Qy   1264  GCCAAGATCGAGGAGCTGCGGCAAGCACCTGCTCGCGCTGGGGCTTCACCACCCCGACAAG  1323
           |||||||||||||||||||||||||||||||||||||||||||||
Db    487  ACCAAGATCGAGGAGCTGAGACAGCAGCATCCTGTTGAGGTGGGACTGACCACCAGACAAG   546

Qy   1324  AAGACCACCAGAAGGAGGAGCCCCCCTTCCTGTGGGATGGGCTACGAGCTGCACCCGACAAGTGG  1383
           |||||||||||||||||||||||||||||||||||||
Db    547  AAGCACCCAGAAGAAGGACCTCCCTTCCTGTGTGATGGGCTACGAACTGCATCCTGACAAGTGG   606

Qy   1384  ACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAG  1443
           ||||||||||||||||||||||||||||||||||||||
Db    607  ACAGTGCAGCCCATCGTGCTCCTGAGAAGGACCTGACTGTGAACGACATACAGAAG   666

Qy   1444  CTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCCAGCTG  1503
           |||||||||||||||||||||||||||||||||||||||
Db    667  CTCGTGGGCAAGTTGAACTGGGCAAGCCAGATCTACCCAGGCATCAAAGTTAGGCAGCTG   726

Qy   1504  TGCAAGCTGCTCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCC  1563
           |||||||||||||||||||||||||||||||||||||||||||
Db    727  TGCAAGCTGCTTCGAGGACCAAGGCCACTGACAAGTGATCCCACTGACAGAGGAAGCA   786

Qy   1564  GAGCTGGAGCTGGCCGAGAACCGGGAGATCCTGCGCGAACGGGCCCGTGCACGGCGTGTACTAC  1623
           |||||||||||||||||||||||||||||||||||||
Db    787  GAGCTAGAACTGGCAGAACCGGACAGAATCCTGAAGGAGCCAGTACATGGAGTGTACTAC   846

Qy   1624  GACCCCAGCAGCAAGGACCTGGTGCCCGAGATCCAGAAGCAGGGCCCACGACCAGACCTAC  1683
           ||||||||||||||||||||||||||||||||||||||||||
Db    847  GACCCAAGCAAGGACCTGATCGCACGGACAGAAGCAGTCCAGAGAGAGGCAAGGCCAATGAGGACCTAC   906

Qy   1684  CAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCGCAAGATGCCCACC  1743
           |||||||||||||||||||||||||||||||||||||||||
Db    907  CAAATCTACCAGGAGCCCTTCAAGAACCTGAAGACAGGCAAGTACGCCGAAGGATGAGGGGT   966

Qy   1744  GCCCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGC  1803
           ||||||||||||||||||||||||||||||||||||
Db    967  GCCCACCAACGATGTGAAGCAGCTGACAGTGAAGCAGATCCACCACCAGAGAGC  1026

Qy   1804  ATCGTGATCTGGGCGCAAGGACCCCCAGTTCCGCTGCCCATCGAGGAGGACCTGGGAG  1863
           |||||||||||||||||||||||||||||||||||||
Db   1027  ATCGTGATCTGGGCGCAAGGACTCCCAAGTTCAAGCTGCCCATACAGAAGGACATGGGAG  1086

Qy   1864  ACCTGGTGACCGACTACTGGCAGGCCACCTGGACCATCTCGGTTCGTGAACACC  1923
           |||||||||||||||||||||||||||||||||||||||
Db   1087  ACATGGTGACCGAGTACTGCGAAGCCACCTGGACCATCCTGAGTTGGAGTTCGTGAACACC  1146

Qy   1924  CCCCCCTGGTGAAGCTGTGTGTACCAGCTGGAGAAGGAGCCCATCATCGCGGCCGAGACC  1983
           |||||||||||||||||||||||||||||||||||||
Db   1147  CCTCCCTTGGTGAAACTGTGGTATCAGCTGGAGGTGGAGAAGGAACCCATCGTGGGAGCAGAGACC  1206

Qy   1984  TTCTACGTGACGGCGGCGCCCAACCGAGCCAAGATCGGCAAGGCCGGCTACGTGACC  2043
           ||||||||||||||||||||||||||||||||||||||
Db   1207  TTCTACGTGGATGGGGACAGCCAGCCAACAGGAGACCAAGCTACGGCAAGGCAAGCTACGTGACC  1266

Qy   2044  GACCGGGCGCGGCCAGAAGATCGTGACGACCACCCCCGACCAACCAGAAGACCGAGCTG  2103
           |||||||||||||||||||||||||||||||||||||
Db   1267  AACCGAGGACAGACACAGAAGTGGTGACCCTGACCGACACAACCAACCAGAAGACTGAGCTG  1326

Qy   2104  CAGGGCCATCCAGCTGGCCCTGCAGGACACCCTGTGCGACGACTGGAGCATCGTGACCGACAGC  2163
           ||||||||||||||||||||||||||||||||||||||
Db   1327  CAAGCCATCCAATCTACCTGCTCTGCAAGACAGCAGCCTGACCGGACTGACCAACGAAGATCGTGACGACTCA  1386

Qy   2164  CAGTACGCCCTGCTGGACCATCATCAGGCCCAGGCCCGACAAGAGCCGAGGCGGTGGTGAAC  2223
           ||||||||||||||||||||||||||||||||||||||
Db   1387  CAGTACGCACTGGGCATCATCAGCCAGCATCCAGCACGGACTGACCATCGTGACGACTGGTGAAC  1446

Qy   2224  CAGATCATCGAGCAGCTGATCGAGGCCCAAGATGATCGGCGGCTTCATCAAGTGTGCGCCCAGT  2283
           |||||||||||||||||||||||||||||||||||||||
Db   1447  CAGATCATCGAGCAGCTGATCGAGGCCGAAGAAGTGTACTGGCCATGGGTACCCAGGACCACAC  1506
```

```
Db  2840  ACCCATCAAAAGACTTAATAGCTGAAATACAGAGAAACAGGGGCATGACCAATGGCACCG  2899
               |||||||||||| ||  ||  ||  ||  ||   ||||| |||  || |||| || |||
Qy  1685  AGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGGCAGTCAAGTACCGCCAAGATGC  1744

Db  2900  AAATTTACCAAGAACCATTCAAAAATCTGAAAACACAGAAAGTATGCAAAAGAGGAGGACTG  2959
               || ||  |||  || || || || |||||   || ||   ||   ||| || |||||
Qy  1745  CCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCGAGAAGATCGCCATGGAGAGCA  1804

Db  2960  CCCACACTAATGACGTAAAACAGTTAACGAGGCAGTGCGTGCAAAAAGATAGCAGCGAAAGCA  3019
              |||||| || |||| || ||||| |  ||||| ||||  ||  |||||| || || ||||
Qy  1805  TCGTGATCTCGGGCAAGACCCCCAAGTTCGGCCTGCCATCCGAAGGAGACCTGGGAGA  1864

Db  3020  TAGTAATATGGGGAAAGACTCCTAAATTAGATTACCCATCCATCCAGAAAGAAACATGGGAAA  3079
               || |||| |||| ||||||| |  || || ||ccc||cc|||||||| |||| |||||
Qy  1865  CCTGGTGGACCGACTACTGGCAGGACCACCTGGAGGACCCCATCATCCGGCCGCCCGAGACCT  1924

Db  3080  CATGGTGGACAGACTATTGGCAAGCCACCTGAGTGTGGGAGTTTGTTAATACCC  3139
              |||||||| |||||||| |||||  ||||| |   || |
Qy  1925  CCCCCCTCGTGAAGCTGTGGTACCAGCTGAGCTGGTGACTTCCGCGAGC  1984

Db  3140  CTCCCCTAGTAAAATTATGGTACCAGCTGAGAAAGAACCCATATTAGGAGCAGAGACTT  3199
               ||||| |||| || ||| |||||||||| |   || |||| |||  ||||| ||
Qy  1985  TCTACGTGGACGGCGCCGCAACCGCCGAGATCGGCAAGGCCGGCTACGTGACCG  2044

Db  3200  TCTATGTAGATGGAGCAGCTAATCGGCAGCTAAACTAAATAGGAAAAGCAGGGTATGTTACTG  3259
              ||| || || |||||||| |   ||||| |    ||| || ||| ||   |||| ||
Qy  2045  ACCGGGCGCCGCAGAAGATCGTGAGCGTGCAGGAGACCGAGCAGAGCAGC  2104

Db  3260  ACAGAGGAAGGCAGAAGAAATTGTTTCTCTAACTGAAACAAAATCAGAGACTGAATTAC  3319
              || ||||  |||||  |||   || || ||| ||||| || |||||||| |||  ||
Qy  2105  AGGCCATCCAGCTGGCCCTGCAGGATCAGGATCCCGATGAACCAGAGACCC  2164

Db  3320  AAGCAATTCAGCTTACAGGATTCAGGATCCCAGCGAGATCTGGTGAACC  3379
              ||||||| |||| |||||  |||||||| || ||| |||  ||||||||
Qy  2165  AGTACGCCCTGGGCCATCATCCAGGCCTGGAAGGGCCAGCAGTCAATC  2224

Db  3380  AGTAATGCATTAGGAATCATTCAGCACAACCAGATAGCAGGGAATCGGCAGCAGTCAATC  3439
              ||| | ||  |||||||||| |||| ||  ||  ||||| || ||||||||||||||||
Qy  2225  AGATCATCGAGCAGCTGATCAAGAAGAGTGTACCTGAGCGTGCCGCCCACA  2284

Db  3440  AAATAATAGACAAGTTAATAAAAAGGAAGAGAGTGCTACCTGTCATGGGTACCAGCACATA  3499
              || || || ||||||| || ||||| ||||||| |  ||| ||| |||| ||||||||
Qy  2285  AGGGCATCGGCGCGAACGAGCAGCAGATCGACAAGCTGGTGTGAGCAAGGTGCAAGGGTGC  2344

Db  3500  AAGGAATTGGAGGAAATGAACAAGTAGTATAAATTAGTAGTGGAATCGAGGAGAGTGC  3559
              ||||||| |  || |  |||   |||| ||| |||| |   ||||||||| ||||||
Qy  2345  TGTTCCTGACGGACATCGAT  2364
              ||| |  | || ||  ||
Db  3560  TGTTTCTAGTAGTGGAATAGAT  3579
```

```
Qy   551  ACGACCAGATCCTGATCGAGATCTGGCGCAAGAGGCCATCGGCACCGTGCTGATCGGCC   610
               |||| |||| ||   ||| | ||  || ||  ||||| || || || |
Db  1760  ATGATCAAATACTTATGAGAAATTTGTGCAACAAAAAGAGGCCATAGTACAGTATTAGTAGGAC  1819

Qy   611  CCACCCCGTGAACATCATCGGCGCCAACATGCTGGCTGACCCTGAACT   670
              ||| | || || |||  || || ||| || |||||||||   |||||
Db  1820  CTACATCTGTCAACATAATTGGAAGAAATATGTTGACCCAGATTGGTTGTACTTTAAACT  1879

Qy   671  TCCCCATCAGCCCCATCGAGACCGTGCCGTGAGCTGAAGCATTAAAGCCAGGGCCCCA   730
              ||||||||||| ||||||||| || |||  |||||||  ||   ||  || || ||||
Db  1880  TTCCAATTAGTCCTATTGAGACTAGCATTAAGAAAGCCCAAGGAATGGCATCTGCAGGG  1939

Qy   731  AGGTGAAGCAGTGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGGCCATCTGCGAGG   790
               |||| |||| || |||| |||  ||||| || |   || || ||| ||||| | ||||
Db  1940  AGGTTAAACAATGGCCATTGACAGAGGAAAAAATAAAAGCATAAGAGAAATTTGTACAG  1999

Qy   791  AGATCGGAGAGGAGGGGCAAGATCACCACCAAGATCGGCGCCCCGAACCCCTACAACCCCG   850
              ||| || || ||  |||||| || ||| ||| |   || |||| || ||  || || |
Db  2000  AAATGGAAAAAGAGGGGAAAATTTCAAAAATCGGGCTGCAAAATCCATATAACACTCCAG  2059

Qy   851  TGTTCGCCATCAGCACCAAGTGCGCAAGCTCGTGTGGACTTCCGCGAGC   910
              || || ||||  |||||||| ||| ||| | ||| | ||||| |||||
Db  2060  TATTTGCCATAAAAAGAAGAAGACAGTACTAAGTGAGAAATTTAGTAGATTTCAGGGAC  2119

Qy   911  TGAACAAGCGCACCCAGGACTTCAGACTTTTGGGAAGTCCAATTAGGGATCCCCGCCTACT  970
              || ||||| |||||| ||||| ||| || |||| |||  || || |||   || | |||
Db  2120  TCAACAAAAGAACTCAAGACTTTGACGTGCTGACGTGGGGCATCCCCCAGATACCCGGCC  2179

Qy   971  TGAAGAGAAGAGAGCGTGACCGTGCACGGTGACTACTGACAGTGGCATAGTCCCCGCCTT  1030
               || || ||||| |||| || || ||  ||||||||| | ||||| ||   ||||| ||
Db  2180  TAAAGAGAAAAATCGAGTGAGCGTGGGGGGATGGCATAATCATATGTTACTTTCTCAGTTCCTT  2239

Qy  1031  TGGACGAGGAGCTTCGCCAAGTACACCACCGCGCCTTCACCATCCCAGCATCAACGAAGACACCC  1090
              || ||| |||  |||||||||| ||||||| || ||||| |||  || ||  || |||| |
Db  2240  TAGATGAGGCTTCAGGAAATATACTGCATTCACCTAGTTCACCACTAGTATAAACAATGAAACAC  2299

Qy  1091  CCGGCATCCGTCGTACCAGTGTACACGTGGCCTGAGAAGGGCTGGAAGGGCCAGCATCT  1150
              || ||||| ||   |||| ||| |||| ||| ||  ||  |||| || |||||||| |
Db  2300  CAGGATTAGGTATCAATATCATTGTGCTTTCCACAGGGGGATGGAAAGGGCACCAGCAGAATAT  2359

Qy  1151  TCCAGAGCAGCATGACCAAGAGTCTGGAGCGCCCTTCCGCGGCGACCTGGAGATCGTGA  1210
              ||||||| |||||||| ||||| || || ||||| || || ||||| |||||| ||||
Db  2360  TCCAAAATAGTATGACAAAAAGTCTTAGGGCACCCTTTAGGGGCACAAAAACCAAGAAATAGTGA  2419

Qy  1211  TCTACCA------GGCCCCCTGTACGTGCAGTGGGCAGCCAACCTGGACGACTGGGCAGCCAACA  1264
              ||||| |       ||||| || || || ||| |||| || || ||  || ||||| ||||||
Db  2420  TCTATCAATATATGGGATGGACTTGTACGTAGATCTGACTTAGAAAATAGGGGCAACATAGAG  2479

Qy  1265  CCAAGATCGAGGAGCTGCGCAAGCTGCGCAAGCCCGCCGACCCCCGACAAGA  1324
              ||||||| |||||||| ||||||| | |||| | ||  | ||| ||||||
Db  2480  CAAAAATAGAGAGAGTTAAGAGACATCATTCTTTGGATGGGATTTACCCAGACAAGA  2539

Qy  1325  AGCACCAGGAAGAGAGCCCCCTTCCTGGATGGGCTACGAGCTGCACGAGCAAGTGGA  1384
              ||||||||||| ||||||| |  || |||||| |||||||||| ||||| ||| ||
Db  2540  AACATCAGAAGAAAGACCTCCATTTCTTCTTTGGATGGGTATGGGTATTACCTGACAAATGGA  2599

Qy  1385  CCGTGCCAGCCATCGAGCTGCCCAGCTGCCCCAGCCTGGCGCCAGATCTGGGCGCCAGT  1444
              || || || | ||| |||| ||||||  || ||||  || ||||||||||  | |||
Db  2600  CAGTACAACCTATACACAGCTGCGAACTGGGATTCATTTACCCGGTTTAAGTGTTAAGTGA  2659

Qy  1445  TGGTGGGCAAGCTGAACTGGCATTGGGACTAAAAACTGGGCGCCCCTGACCGAGGAGGCCG  1504
              ||||||||| | ||||| ||||  |||  | ||||| ||||||||| ||||| || |  |
Db  2660  TAGTGGAAAAATTAAACTGGGACTGGGATCAGTCGATTTAACCAGAGTTAACTGACGAAGAGC  2719

Qy  1505  GCAAGCTGCTGCGCGCGCCCAAGCCCGCCGAGATCCATGTGCCCCTGACCGAGGAGGCCG  1564
              |||| || || |||| |||| ||||  ||| ||  | |||  ||  ||||||||| || |
Db  2720  GTAAACTCCTTAGGGGACCTGGGGATGGATGGAGACCATTAACAGAGGACTAACTGAAGAGCAG  2779

Qy  1565  AGCTGGAGCTGGCCGAGAGTGGCCGAGACCGTGCCGGCGTGTACTACG  1624
              ||||||||||| | ||  || |||| || ||||| |||  |||   |
Db  2780  AATTAGAATTGGCAGAAATTTTAAAAAGGAAAACCAGTACATGGAGTATATTATG  2839

Qy  1625  ACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGCGCCACGAGGACCAGACCTACC  1684
```

```
          Copyright  GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      April 10, 2004, 12:49:38 ; Search time 593.333 Seconds
                                          (without alignments)
                             15572.780 Million cell updates/sec

Title:            US-09-610-313-31
Perfect score:    2463
Sequence:         1 gtcgacgccaccatggccga.........gggctagcaccggtgaattc 2463

Scoring table:    IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched:         2475585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters:     4951170

Minimum DB seq length:  0
Maximum DB seq length:  2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

                                SUMMARIES

Database :   Published Applications_NA:*
        1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
        2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
        3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
        4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
        5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
        6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
        7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
        8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
        9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
       10:  /cgn1_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
       11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
       12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
       13:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
       14:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
       15:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
       16:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
       17:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
       18:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
Result         %
No.    Score   Query   Length  DB  ID                      Description
               Match
----------------------------------------------------------------------------
  1  2455.4   99.7    2457    14  US-10-190-435-44     Sequence 44, Appl
  2  2455.4   99.7    2457    14  US-10-190-305A-38    Sequence 38, Appl
  3  2430.2   98.7    2457    14  US-10-190-435-45     Sequence 45, Appl
  4  2430.2   98.7    2457    14  US-10-190-305A-39    Sequence 39, Appl
  5  2422.6   98.4    2445    14  US-10-190-435-43     Sequence 43, Appl
  6  2422.6   98.4    2445    14  US-10-190-305A-37    Sequence 37, Appl
  7  2415.6   98.1    3930    14  US-10-190-435-9      Sequence 9, Appli
  8  2414     98.0    3930    14  US-10-190-435-10     Sequence 10, Appl
  9  2414     98.0    3930    14  US-10-190-435-11     Sequence 11, Appl
 10  2414     98.0    5184    14  US-10-190-435-58     Sequence 58, Appl
 11  2414     98.0    5184    14  US-10-190-305A-82    Sequence 82, Appl
 12  2383.6   96.8    3531    14  US-10-190-435-13     Sequence 13, Appl
 13  2382     96.7    3537    14  US-10-190-435-14     Sequence 14, Appl
 14  2382     96.7    3537    14  US-10-190-435-15     Sequence 15, Appl
 15  2381     96.7    5145    14  US-10-190-435-12     Sequence 12, Appl
```

```
 16  2381     96.7    5145    14  US-10-190-305A-12    Sequence 12, Appl
 17  2322.6   94.3    3607    14  US-10-190-435-48     Sequence 48, Appl
 18  2322.6   94.3    3607    14  US-10-190-305A-42    Sequence 42, Appl
 19  2322.6   94.3    3624    14  US-10-190-435-47     Sequence 47, Appl
 20  2322.6   94.3    3624    14  US-10-190-305A-41    Sequence 41, Appl
 21  2304.4   93.6    3597    14  US-10-190-435-46     Sequence 46, Appl
 22  2304.4   93.6    3597    14  US-10-190-305A-40    Sequence 40, Appl
 23  2142     87.0    2466    14  US-10-241-009-31     Sequence 31, Appl
 24  2142     87.0    2466    14  US-10-190-434B-31    Sequence 31, Appl
 25  2142     87.0    2466    14  US-10-190-305A-33    Sequence 33, Appl
 26  2122.8   86.2    2472    14  US-10-241-009-32     Sequence 32, Appl
 27  2122.8   86.2    2472    14  US-10-190-434B-32    Sequence 32, Appl
 28  2122.8   86.2    2472    14  US-10-190-305A-34    Sequence 34, Appl
 29  2115.2   85.9    2460    14  US-10-190-434B-30    Sequence 30, Appl
 30  2115.2   85.9    2460    14  US-10-190-305A-30    Sequence 30, Appl
 31  2115.2   85.9    2460    14  US-10-190-305A-32    Sequence 32, Appl
 32  2114.2   85.8    3564    14  US-10-241-009-13     Sequence 13, Appl
 33  2114.2   85.8    3564    14  US-10-241-009-14     Sequence 14, Appl
 34  2114.2   85.8    3564    14  US-10-190-434B-13    Sequence 13, Appl
 35  2114.2   85.8    3564    14  US-10-190-434B-14    Sequence 14, Appl
 36  2113.6   85.8    4716    14  US-10-190-435-17     Sequence 17, Appl
 37  2113.6   85.8    4716    14  US-10-190-305A-13    Sequence 13, Appl
 38  2109.8   85.7    3999    14  US-10-241-009-9      Sequence 9, Appli
 39  2109.8   85.7    3999    14  US-10-190-434B-9     Sequence 9, Appli
 40  2108.2   85.6    3999    14  US-10-241-009-10     Sequence 10, Appl
 41  2108.2   85.6    3999    14  US-10-241-009-11     Sequence 11, Appl
 42  2108.2   85.6    3999    14  US-10-190-434B-10    Sequence 10, Appl
 43  2108.2   85.6    3999    14  US-10-190-434B-11    Sequence 11, Appl
 44  2108.2   85.6    5283    14  US-10-241-009-54     Sequence 54, Appl
 45  2108.2   85.6    5283    14  US-10-190-434B-54    Sequence 54, Appl
```

ALIGNMENTS

```
RESULT 1
US-10-190-435-44
; Sequence 44, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
;  LENGTH: 2457
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: p2Polopt.YM_C
US-10-190-435-44

Query Match             99.7%;  Score 2455.4;  DB 14;  Length 2457;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 2456;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1  GTCGACGCGCACCATGGCCGAGGCCACGGCCACCAGCGCCACCATCCTGATCCAG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  GTCGACGCGCACCATGGCCGAGGCCACGGCCACCAGCGCCACCATCCTGATCCAG  60

Qy      61  CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  CGCAGCAACTTCAAGGGCCCAAGCGGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGC  120

Qy     121  CACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCGTGCTGGAGTGCGGCAAGGAG  180
```

```
Db  1201 GAGATCGTGATCTACCAGGCCCCCTGTACGTGGCAGCGACCTTGGAGATCGGCCAGCGAC 1260

Qy  1261 CGGCCCAAGATCGAGGAGTCGCAAGCAACGCACCTGCTGCGCGCTTCACCACCCCCGACAG 1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CGGCCCAAGATCGAGGAGTCGCAAGCAACGCACCTGCTGCGCGCTTCACCACCCCCGACAG 1320

Qy  1321 AAGAAGCACCAGAAGGAGAGCCCCCTTCCTGTGGATGGGTACGAGCTGCACCCCGACAAG 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 AAGAAGCACCAGAAGGAGAGCCCCCTTCCTGTGGATGGGTACGAGCTGCACCCCGACAAG 1380

Qy  1381 TGGACCGTGCAGCCATCGAGCTGCCCGAGAAGAGAGAGCTTGACCGTGAACGACATCCAG 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TGGACCGTGCAGCCAGCCATCGAGCTGCCCGAGAAGAGAGAGCTGGACCGTGAACGACATCCAG 1440

Qy  1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGATCTACCCGGCATCAAGGTCGCCAG 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGATCTACCCGGCATCAAGGTGCGCCAG 1500

Qy  1501 CTGTGCAAGCTGCTGCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CTGTGCAAGCTGCTGCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy  1561 GCCGAGCTGGAGCTGCCGAGAACCGGAGAATCCTGCGCGAGGACGGCGTGACGGGCGTGTAC 1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GCCGAGCTGGAGCTGCCGAGAACCGGAGAATCCTGCGCGAGGACGGCGTGACGGGCGTGTAC 1620

Qy  1621 TACGACCCCAGCAAGGACCTGGTGCCCGAGATCGTGAAGACCGAAGAGGACCACGACCAGTGGACC 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 TACGACCCCAGCAAGGACCTGGTGCCCGAGATCGTGAAGACCGAAGAGGACCACGACCAGTGGACC 1680

Qy  1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGGTACGCCAAGATGCGC 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGGTACGCCAAGATGCGC 1740

Qy  1741 ACCGCCCACACCAACGACGTGAAGCAGTGTGACGAGCCCGTGCCGTGCCATGGAG 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ACCGCCCACACCAACGACGTGAAGCAGTGTGACGAGCCCGTGCCGTGCCATGGAG 1800

Qy  1801 AGCATCGTGATCTGGGCCAAGACCCCCAAGTTCCGCGCTGCCATCCGAGTTCGTGAAC 1860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGCATCGTGATCTGGGCCAAGACCCCCAAGTTCCGCGCTGCCATCCGAGTTCGTGAAC 1860

Qy  1861 GAGACCTGGTGACCTACTGGCAGGCCACCTGGTGGAGTTCGTGAAC 1920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 GAGACCTGGTGACCTACTGGCAGGCCACCTGGTGGAGTTCGTGAAC 1920

Qy  1921 ACCCCCCCTGTGAAGCTGTGGTACCAGTGTGGAGAAGGAGCCCATCATCGGCGGCCGAG 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ACCCCCCCTGTGAAGCTGTGGTACCAGTGTGGAGAAGGAGCCCATCATCGGCGGCCGAG 1980

Qy  1981 ACCTTCACGTGACGGCCGGCCGCCAACCGGGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ACCTTCACGTGACGGCCGGCCGCCAACCGGGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040

Qy  2041 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACGGCAGCGAGGTGAACATCGTGACCGAC 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACGGCAGCGAGGTGAACATCGTGACCGAC 2100

Qy  2101 CTGCAGGCCATCCAGCTGGCCCTGACCGAGGCAACGACCAGCGCAGCGAGGTGAACATCGTGACCGAC 2160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGCAGGCCATCCAGCTGGCCCTGACCGAGGCAACGACCAGCGCAGCGAGGTGAACATCGTGACCGAC 2160

Qy  2161 AGCCAGTACGCCCTGGCCATCATCCGGCATCATCCAGGCTGAAGGGCAGCGAGCTGGTG 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 AGCCAGTACGCCCTGGCCATCATCCGGCATCATCCAGGCTGAAGGGCAGCGAGCTGGTG 2220

Qy  2221 AACCAGATCATCGAGCAGCTGATCAAGAAGAGAAGTGTACCTGAGCTGGTGCCCGCC 2280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AACCAGATCATCGAGCAGCTGATCAAGAAGAGAAGTGTACCTGAGCTGGTGCCCGCC 2280

Qy  2281 CACAAGGGCATCGGCGCAACGACCAGCAGATCGACAAGCTGGTGAGCAAGCTGTGAGCATCCGCAAG 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACAAGGGCATCGGCGCAACGACCAGCAGATCGACAAGCTGGTGAGCAAGCTGTGAGCATCCGCAAG 2340
```

```
Db  121  CACATCGCCGCAACTGCCGCGCCCCGCCCCAAGAAGAAGGGCTGCTGGAGTGCGGCAAGGAG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  181  GGCCACCAGATGAAGGACTGCACCGAGAGCTTCTTCCGCGAGGACTGGCC 240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GGCCACCAGATGAAGGACTGCACCGAGAGCTTCTTCCGCGAGGACTGGCC 240

Qy  241  TTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCGACGACCGCCAACAGCCCCACC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCGACGACCGCCAACAGCCCCACC 300

Qy  301  AGCCGCGAGCTGCAGGTGCGCGGCGACACCCCGAGCGAGGCCGGCGCGCCGAGCGCCAG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  AGCCGCGAGCTGCAGGTGCGCGGCGACACCCCGAGCGAGGCCGGCGCGCCGAGCGCCAG 360

Qy  361  GGCACCCTGAACTTCCCCAGATCACCCTGTGCAGGCCCCCTGTGAGCATCAAGGTG 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GGCACCCTGAACTTCCCCAGATCACCCTGTGCAGGCCCCCTGTGAGCATCAAGGTG 420

Qy  421  GGGGGCCAGTACGACAAGGAGGCCCTGCTGGACACCGGCCAACGACCCGTCTGGAGGAG 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GGGGGCCAGTACGACAAGGAGGCCCTGCTGGACACCGGCCAACGACCCGTCTGGAGGAG 480

Qy  481  ATGAGCCTGCCCGGCCAGTGGAAGCCCAAGATGATCGGCGCAAGTCGGCGCTTCATCAAG 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ATGAGCCTGCCCGGCCAGTGGAAGCCCAAGATGATCGGCGCAAGTCGGCGCTTCATCAAG 540

Qy  541  GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCCACCGTG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCCACCGTG 600

Qy  601  CTGATCGGCCCACCCCGTGAACATCATCGGCCGACCCAGTGGCTGC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CTGATCGGCCCACCCCGTGAACATCATCGGCCGACCCAGTGGCTGC 660

Qy  661  ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGAAGCTGAAGCCCGGCATG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGAAGCTGAAGCCCGGCATG 720

Qy  721  GACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGATCAAGGCCCTGACCGC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  GACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGATCAAGGCCCTGACCGC 780

Qy  781  ATCTGGAGGAGATGGAGAAGGGCAAGATCACCAGGATCGGCCCGAGAACCCCTAC 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  ATCTGGAGGAGATGGAGAAGGGCAAGATCACCAGGATCGGCCCGAGAACCCCTAC 840

Qy  841  AACACCCCGTGTTCGCCATCAAGAAGAAGACAGCACCAAGTGGCGCAAGTCGTGGAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  AACACCCCGTGTTCGCCATCAAGAAGAAGACAGCACCAAGTGGCGCAAGTCGTGGAC 900

Qy  901  TTCCGGAGCTGAACAAGCGCACCCAGGACTTCGGCGAGGTGCAGCTGCAGCTGGGCATCCCCAC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TTCCGGAGCTGAACAAGCGCACCCAGGACTTCGGCGAGGTGCAGCTGCAGCTGGGCATCCCCAC 960

Qy  961  CCCGCCGGCCTGAAGACCGACCTGGACGTGGGCGACGCCCTACTTC 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  CCCGCCGGCCTGAAGACCGACCTGGACGTGGGCGACGCCCTACTTC 1020

Qy  1021 AGCGGTGCCCCTGACGAGACTTCCGCAGAGACTTCACCGACATCAAC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AGCGGTGCCCCTGACGAGACTTCCGCAGAGACTTCACCGACATCAAC 1080

Qy  1081 AACGAGACCCCGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGAAGGGCAGC 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 AACGAGACCCCGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGAAGGGCAGC 1140

Qy  1141 CCCAGCATCTTCAGAGCAGCATGACCAAGATCCTGAGCCCTTCCGCGCCCAACCCC 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CCCAGCATCTTCAGAGCAGCATGACCAAGATCCTGAGCCCTTCCGCGCCCAACCCC 1200

Qy  1201 GAGATCGTGATCTACCAGGCCCCCTGTACGTGGCAGCGACCTGAGATCGGCCAGCGAC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   2341 GTGCTGTTCCTGGACGGCGGCATCGATGGCGGCATCGTGATCGATCTGATCTACCAGTACATGGACGACCTG 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCGGCATCGATGGCGGCATCGTGATCGATCTGATCTACCAGTACATGGACGACCTG 2400

Qy   2401 TACGTGGGACGAGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGTTAGCACCGGT 2457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGGACGAGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGTTAGCACCGGT 2457
```

RESULT 2
US-10-190-305A-38
; Sequence 38, Application US/10190305A
; Publication No. US20030196621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt.YM_C
US-10-190-305A-38

Query Match   99.7%; Score 2455.4; DB 14; Length 2457;
Best Local Similarity 100.0%; Pred. No.0;
Matches 2456; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 GTCGACGCGCCACCATGGCCGAGGCCATGAGCGCCAGGCCGCCAACATCCTGATGCAG   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTCGACGCGCCACCATGGCCGAGGCCATGAGCGCCAGGCCGCCAACATCCTGATGCAG   60

Qy   61 CGCAGCAACTTCAAGGGCCCCGCCCGCCAAGTGCCATCATCAAGTGCTTCAACTGCGGCAAGAGGGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGCAGCAACTTCAAGGGCCCCGCCCGCCAAGTGCCATCATCAAGTGCTTCAACTGCGGCAAGAGGGGC 120

Qy   121 CACATCGCCCGCAACTGCCGCGCCCCCCGCCCCCCCCAAGAAGGGCGTGCTTGGAAGTGCGGCAAGGAG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACATCGCCCGCAACTGCCGCGCCCCCCGCCCCCCCCAAGAAGGGCGTGCTTGGAAGTGCGGCAAGGAG 180

Qy   181 GGCCACCAGATGAGGACTGCACCGAGGGCCAGGCCAACTTCTTCGGGAGGACCTGGCC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCCACCAGATGAGGACTGCACCGAGGGCCAGGCCAACTTCTTCGGGAGGACCTGGCC 240

Qy   241 TTCCCCAGGGCAAGGCCCGCGAGTTCCCAGCCAGGAGAACGCGGCCAACAGCCCCACC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCCCAGGGCAAGGCCCGCGAGTTCCCAGCCAGGAGAACGCGGCCAACAGCCCCACC 300

Qy   301 AGCGCGGAGCTGCAGGTGCGGGCGGCGACAACCCCCAGCGGCCGAGGCCGCCAGCGCCAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCGCGGAGCTGCAGGTGCGGGCGGCGACAACCCCCAGCGGCCGAGGCCGCCAGCGCCAG 360

Qy   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG 420

Qy   421 GGCGGCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG 480

Qy   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540
```

```
Qy    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600

Qy    601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC 660

Qy    661 ACCCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720

Qy    721 GACGGCCCCAAGTGAAGCAGTGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACGGCCCCAAGTGAAGCAGTGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780

Qy    781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840

Qy    841 AACACCCCGTTCGCCATCAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACACCCCGTTCGCCATCAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900

Qy    901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC 960

Qy    961 CCCGCGGCCTGAAGAAGAAGAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCCGCGGCCTGAAGAAGAAGAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTC 1020

Qy   1021 AGGGTGCCCCTGGACGAGGACTTCCGGCAAGTACACCCGCCTTCACCATCCCCAGCATCAAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGGGTGCCCCTGGACGAGGACTTCCGGCAAGTACACCCGCCTTCACCATCCCCAGCATCAAC 1080

Qy   1081 AACGAGACCCCGGCATCCCGGAGGCCATGAGCAGGCCAACATCCTGATGCAG 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACGAGACCCCGGCATCCCGGAGGCCATGAGCAGGCCAACATCCTGATGCAG 1140

Qy   1141 CCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCCGCAACCCC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCCGCAACCCC 1200

Qy   1201 GAGATCGTGATCTACCAGGCCCCCCGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCAGGCCCCCCGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1260

Qy   1261 CGCGCCAAGATCGAGGAGCTGCGCCAAGCTTCTTCGCGGAGGACTTCACCCCCGGCC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGCCAAGATCGAGGAGCTGCGCCAAGCTTCTTCGCGGAGGACTTCACCCCCGGAC 1320

Qy   1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGTGGATGGGCTAGCAGCTGCACCCCGACAAG 1380

Qy   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGAGCTGGACCGTGAAGGACATCCAG 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGAGCTGGACCGTGAAGGACATCCAG 1440

Qy   1441 AAGTGGTGGCAAGCTGAACTGGGCCAGATCACCCGGATCAAGGTGCGCCAG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGTGGTGGCAAGCTGAACTGGGCCAGATCACCCGGATCAAGGTGCGCCAG 1500

Qy   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGAG 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGAGGTGATCCCCCTGACCGAGGAG 1560

Qy   1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCGTGGGCGAGCCGTGCACCGGCGTGTAC 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCGTGGGCGAGCCGTGCACCGGCGTGTAC 1620

Qy   1621 TACGACCCCAGCAAGGACCTGGTTGCCGAGATCCAGAAGCAGGCACCAGACCAGTGGACC 1680
```

; SEQ ID NO 45
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt_C
US-10-190-435-45

Query Match   98.7%; Score 2430.2; DB 14; Length 2457;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2448; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
Qy      7 GCCACCATGGCCCAGGCCATGAGCCAGGCCACCATCCTCGATGCAGCGCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCCAGGCCATGAGCCAGGCCACCATCCTCGATGCAGCGCAGC 60

Qy     67 AACTTCAAGGGCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 120

Qy    127 GCCGGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCAC 186
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCGGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCAC 180

Qy    187 CAGATGAAGGACTGCACCGAGCGGCCAACTTCTTCCGCGAGCAGCTGGCCTTCCCCC 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAAGGACTGCACCGAGCGGCCAACTTCTTCCGCGAGCAGCTGGCCTTCCCCC 240

Qy    247 CAGGGCAAGGCCCCGGAGTTCCCCAGCGAGCAGAACCGGCCAACAGCCCCACCAGCCGC 306
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCAAGGCCCCGGAGTTCCCCAGCGAGCAGAACCGGCCAACAGCCCCACCAGCCGC 300

Qy    307 GAGCTGCAGGTGCGGCGGCGACAACCCCTGTGGCAGCCCCTGGTGCATCAAGGTGGGCGGC 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGCTGCAGGTGCGGCGGCGACAACCCCTGTGGCAGCCCCTGGTGCATCAAGGTGGGCGGC 360

Qy    367 CTGAACTTCCCCCAGATCACCCTGTGGCAGCCCCTGGTGCATCAAGGTGGGCGGC 426
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCCAGATCACCCTGTGGCAGCCCCTGGTGCATCAAGGTGGGCGGC 420

Qy    427 CAGATCAAGGAGCCCGTGCTGCAGAGGAAGGTGTACCTGGAGCTGCTGGAGGAGGATGAGC 486
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGCCCGTGCTGCAGAGGAAGGTGTACCTGGAGCTGCTGGAGGAGGATGAGC 480

Qy    487 CTGCCCGGCAAGTGGAAGCCCAAGATCGGCGGCATCGGGCTTCATCAAGGTGCGC 546
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGCCCGGCAAGTGGAAGCCCAAGATCGGCGGCATCGGGCTTCATCAAGGTGCGC 540

Qy    547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCCAACATGCTGCGCCAGTGCTGCTGATC 606
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCCAACATGCTGCGCCAGTGCTGCTGATC 600

Qy    607 GGCCCCACCCCGTGAACATCAGCCCCATCGAACGCCTGCCCGTGAAGCTGAAGCCGCACCCTG 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCCCACCCCGTGAACATCAGCCCCATCGAACGCCTGCCCGTGAAGCTGAAGCCGCACCCTG 660

Qy    667 AACTTCCCCATCAGCCCCATCGAACGCCTGCCCGTGAAGCTGAAGGCCCTGACCGGCCATCTGC 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AACTTCCCCATCAGCCCCATCGAACGCCTGCCCGTGAAGCTGAAGGCCCTGACCGGCCATCTGC 720

Qy    727 CCCAAGGTGAAGCAGTGCCCTGACCGGAGGAGAAGATCAAGGCCCTGACCGGCCATCTGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCAAGGTGAAGCAGTGCCCTGACCGGAGGAGAAGATCAAGGCCCTGACCGGCCATCTGC 780

Qy    787 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAAACACC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACC 840

Qy    847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 906
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 900

Qy    907 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 966
```

```
Db   1621 TACGACCCCAGCAGGACCAAGGACCTGGTGCCGGAGATCCAGAGCAGGCAGCAGCCAGTGACC 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCCGGAAGTACGGCCAAGATGCGC 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCCGGAAGTACGGCCAAGATGCGC 1740

Qy   1741 ACCGCCCACCACCAACGACGACTGAAGCACCTGAAGCAGAGCCGTGCAGAAGATCGCCATGGAG 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACCGCCCACCACCAACGACGACTGAAGCACCTGAAGCAGAGCCGTGCAGAAGATCGCCATGGAG 1800

Qy   1801 AGGATCGTGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCTGCCTCAGAAGGAGGACCTGG 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCTGCCTCAGAAGGAGACCTGG 1860

Qy   1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920

Qy   1921 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGGCCGAG 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGGCCGAG 1980

Qy   1981 ACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040

Qy   2041 ACCGACCGGGGCCGGCAGAAGATCGTGACCCTGACCGAGACCACCAACCAGAAGACCGAG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACCGGGGCCGGCCCGGCAGAGAAGATCGTGACCCTGACCGAGACCACCAACCAGACCGAG 2100

Qy   2101 CTGCAGGCCATCCAGCTGCCCTGCAGGACACCGGCAGCGAGTGAACATCGTGACCGAC 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGCCATCCAGCTGCCCTGCAGGACACCGGCAGCGAGTGAACATCGTGACCGAC 2160

Qy   2161 AGCCAGTACGCCCCTGGGACGCCCCTGCAGCCCGACAAGAGCGAGAGCCAGCGGTGGTG 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACGCCCCTGGGACGCCCCTGCAGCCCGACAAGAGCGAGAGCCAGCGGTGGTG 2220

Qy   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2280

Qy   2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCGAGGGCATCCCCAAG 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCGAGGGCATCCGCAAG 2340

Qy   2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTG 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTG 2400

Qy   2401 TACGTGGGCAGCGACGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGGCAGCGACGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
```

RESULT 3
US-10-190-435-45
; Sequence 45, Application US/10190435
; Publication No. US20030143248A1
GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0

RESULT 4
US-10-190-305A-39
; Sequence 39, Application US/10190305A
; Publication No. US2003019862 1A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt_C
US-10-190-305A-39

Query Match   98.7%;   Score 2430.2;   DB 14;   Length 2457;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 2448;   Conservative 0;   Mismatches 3;   Indels 6;   Gaps 1;

Qy   1321 AAGAAGACCAGAAGGAGCCCCCTTCCTGTGGATCGGCTACGAGCTGCACCCGACAAG 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGAAGACCAGAAGGAGCCCCCTTCCTGTGGATCGGCTACGAGCTGCACCCGACAAG 1380

Qy   1381 TGGACCGTGCAGCCATCGAGCTGCCCGAGAAGGAGAGTGACCGTGAACGACATCCAG 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGTGACCGTGAACGACATCCAG 1440

Qy   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGGCCATCAAGGTGCGCGCCAG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGGCCATCAAGGTGCGCGCCAG 1500

Qy   1501 CTGTGCAAGCTGCTGCGCGGGGCCAAGGCCCTGACCGTGACCGAGGAG 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGCGGGGCCAAGGCCCTGACCGTGACCGAGGAG 1560

Qy   1561 GCCGAGCTGGAGCTGCCGAGAACCGGCAGAGTCCTGCGCGAGAGTCGTGTAC 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGCTGGAGCTGCCGAGAACCGGCAGAGTCCTGCGCGAGAGTCGTGTAC 1620

Qy   1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAGACCAGCAAGATGCGC 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAGACCAGCAAGATGCGC 1680

Qy   1681 TACCAGATCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGACGATCGGC 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACCAGATCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGACGATCGGC 1740

Qy   1741 ACGGCCCACCACCAACGACGTGAAGCAGCTGACCGTGCGAGAAGATCGCCATGGAG 1800
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACGGCCCACCACCAACGACGTGAAGCAGCTGACCGTGCGAGAAGATCGCCATGGAG 1800

Qy   1801 AGGATCGTGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860

Qy   1861 GAGACCTGGTGGACCGACTACTGGCAGGGCCACCTGGAGTTCGTGAAC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGACCTGGTGGACCGACTACTGGCAGGGCCACCTGGAGTTCGTGAAC 1920

Qy   1921 ACCCCCCCGTGTGAAGCTGTGTGTACCAGCTGGAGGAGCCCATCGTGAACATCGTGACCGAG 1980
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 ACCCCCCCGTGTGGTGAAGCTGTGTACCAGCTGGAGCAGCCCCATCATCGTGAACATCGTGACCGAG 1980

Qy   1981 ACCTTCTACGTGGACGGGCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACCTTCTACGTGGACGGGCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040

Qy   2041 ACCGACGGGGCGGCGCCGAGAGATCGTGAGGCTGACCGAGACGAGACCGAG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 ACCGACGGGGCCGGCGCCGCCAGAGATCGTGAGGCTGACCGAGACCAGAGACCGAG 2100

Qy   2101 CTGCAGGCCATCCAGCTGCCCTGAGGGGCCATCAACAACGAG 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGCAGGCCATCCAGCTGCCCTGAGGGGCATCAACAACGAG 2160

Qy   2161 AGCCAGTACGCCCTCGGGCATCAGGGCCAGCGGAGGGCAGCCCCAGC 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCCAGTACGCCCTCGGGCATCAGGGCCAGCGGAGGGCAGCCCCAGC 2220

Qy   2221 AACCAGATCATCGAGCAGCTGATCGAAGAGAAGCAAGAGGTGTACCTGGTGTACCTGGTCCCGCC 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AACCAGATCATCGAGCAGCTGATCGAAGAGAAGCAAGAGGTGTACCTGGTGTACCTGGTCCCGCC 2280

Qy   2281 CACAAGGGCATCGGCCAACGAGCAGATCGACAAGCTGGTGGAGCAAGGACCGAG 2340
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CACAAGGGCATCGGCCAACGAGCAGATCGACAAGCTGGTGGAGCAAGGACCGAG 2340

Qy   2341 GTGCTGTTCCTGACGGCATGGGGCATCGTGATCTACCAGTACATGACGACCTG 2400
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGACGGCATGGGGCATCGTGATCTACCAGTACATGACGACCTG 2400

Qy   2401 TACGTGGCAGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGTAGACCGGT 2457

Qy   247 CAGGGCAAGGCCCGCGAGTTCCCAGCGAGCAGAACCGGCCAACAGCCCACCAGCCGC 306
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGGGCAAGGCCCGCGAGTTCCCAGCGAGCAGAACCGGCCAACAGCCCACCAGCCGC 300

Qy   307 GAGCTGCAGGTGCGCGGCGACAACCCCCAGCGAGGCCGGCGCCAGGCGCACC 366
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAGCTGCAGGTGCGCGGCGACAACCCCCGACGAGGCCGGCGCCAGGGCACC 360

Qy   367 CTGAACTTCCCCAGATCGCATCCTGTGCAGCGCCCCCTGTGAGCATCAAGGTGGGCGC 426
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTGAACTTCCCCAGATCCGATCCTGTGCAGCGCCCCTGTGAGCATCAAGGTGGGCGC 420

Qy   427 CAGATCAAGGAGGCCCTGCTGGACACCGGCGCGCCGACGCACCCGTGCTGGAGGAGATGAGC 486
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CAGATCAAGGAGGCCCTGCTGGACACCGGCGCGCCGACGCACCCGTGCTGGAGGAGATGAGC 480

Qy   487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGCTTCATCAAGGTGCGC 546
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGCCCGGCAAGTGGAAGCCCAAGATCGGCGGCATCGGCGCTTCATCAAGGTGCGC 540

Qy   547 CAGTACGACACCAGATCCTGATCGGGCCAAGAAGGCCATCGGCACCGTGCTGATC 606
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGTACGACCAGATCCTGATCGGGCCAAGAAGGGCCATCGGCACCGTGCTGATC 600

Qy   607 GGCCCCACCCCGTGAACACATGCTGCGGCTGGGCTGCACCCTG 666
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCCCCACCCCCGTGAACACATGCTGACGGCTGGGCTGCACCCTG 660

Qy   667 AACTTCCCCATCGCCGTGAGACATCATCGAGCCCGTGAAGCTGAAGCCGCATGGACGGC 726
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AACTTCCCCATCGCCGTGAGACATCATCGAGCCCGTGAAGCTGAAGCCGCATGGACGGC 720

Qy   727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGAAGATCAAGGCCCGCCCCGAGAGAGTCTGC 786
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGAAGATCAAGGCCCGCCCCGAGAGAGTCTGC 780

Qy   787 GAGGAGATGGACAGTGGCCCCTGGGAGGTGCAGCTTCTGGGAGAAGGCCGGCATCTCCACACACC 846
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GAGGAGATGGACAGTGGCCCCTGGGAGGTGCAGCTTCTGGGAGAAGGCCGGCATCTCCACACACC 840

Qy   847 CCCGTGTTCGCCATCGAAGAGAAGGACAGCAGCAGCACCAAGTGGCCGCAAGTGGCGCAAGTGGTGGACTTCCGC 906
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CCCGTGTTCGCCATCGAAGAGAAGGACAGCAGCAGCAAGTGGCCGCAAGTGGTGGACTTCCGC 900

Qy   907 GAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCCGCC 966
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCCGCC 960

Qy   967 GGGCTGAAGAAGGAGAGAGCGTGCTGACAGTGGCGGTGGGGCACGGCCTACTTCAGCGTG 1026
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGGCTGAAGAAGGAGAAGAGAGCGTGCTGACAGTGGCGGTGGGGCACGGCCTACTTCAGCGTG 1020

Qy   1027 CCCCTGGACGCAGGACTTCCGCAAGTACACCGGCCTTCACCCATCCCAGCATCAACAACGAG 1086
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGCAGGGACTTCCGCAAGTACACACCGGCCTTCACCCATCCCAGCATCAACAACGAG 1080

Qy   1087 ACCCCCGGCATCGGCATCAGCTACAGTGCTGCGCTGCTCGTGCTGCGCCCCAGC 1146
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCCCCGGCATCGGCATCAGCTACAGTGCTGCGCTGCTGGGAGGGCAGCCCCAGC 1140

Qy   1147 ATCTTCCAGAGCAGCATGACCAGATCCTGGGAGGTCAGCTGGGCATCCCCGAGATC 1206
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCTTCCAGAGCAGCATGACCAGATCCTGGGAGGTGCAGCTGGGCATCCCCGAGATC 1200

Qy   1207 GTGATCTACCA------GGCCCCCGTACGTGGACGTGGACGCTGTGACGTGGGCGACGTGGGCGACCCTGACCCCATCCAGCGCCAGCAC 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGATCTACCAGTCTACATGGACGTGGACGCTGTACGTGGACGCTGTACGTGGGCGACGTGGGCGACCCCCTGACCCCATCCAGCGCCAGCAC 1260

Qy   1261 CGCGCCAAGATCGAGGAGATGCTGCGCGTCGCGCCAAGCACCTGCGGCTTCACCCCCGAC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGCCAAGATCGAGGAGATGCTGCGCGTCGCGCCAAGCACCTGCGGCTTCACCACCCCCGAC 1320

```
Db  2401  TACGTGGGCAGCGGCGGCCCCTAGGATCGATGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 5
US-10-190-435-43
; Sequence 43, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Pol.opt.YMWM_C
US-10-190-435-43

```
Query Match     98.4%;    Score 2422.6;   DB 14;   Length 2445;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 2441;   Conservative 0;   Mismatches 4;   Indels 6;   Gaps 1;
```

```
Qy    7  GCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGC   66
         |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db    1  GCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCTGATGCAGCGCCAGC   60

Qy   67  AACTTCAAGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC  126
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AACTTCAAGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC  120

Qy  127  GCCCGCAACTGCGGCGCCCCCGAGTTCCCGGCGAGCAGGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC  186
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GCCCGCAACTGCGGCGCCCCCGAGTTCCCGGCGAGCAGGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC  180

Qy  187  CAGATGAAGGACTGACCCAGGAGCGGCCAGCCAACTTCTTCCGCGAGAACCTGGCCTTCCCC  246
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CAGATGAAGGACTGACCCAGGAGCGGCCAGCCAACTTCTTCCGCGAGAACCTGGCCTTCCCC  240

Qy  247  CAGGGCAAGCCCCGGAGTTCCCCAGCAGCAGAACCGCGCGAACCGGCCCCACCAGCCGC  306
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  CAGGGCAAGCCCCGGAGTTCCCCAGCAGCAGAACCGCGCGAACCGGCCCCACCAGCCGC  300

Qy  307  GAGCTGCAGTGCGGCGGCCACACCCCCGGCCGGAGGCCGGAGCGCCAGGGCACC  366
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GAGCTGCAGTGCGGCGGCGGCGACAACCCCGGCCGGAGGCCGGAGCGCCAGGGCACC  360

Qy  367  CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGC  426
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGC  420

Qy  427  CAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGC  486
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  CAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGC  480

Qy  487  CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGGGCTTCATCAAGGTGCGC  546
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  CTGGACGCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGGGCTTCATCAAGGTGCGC  540

Qy  547  CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC  606
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC  600
```

```
Qy  607  GGCCCCACCCCCGTGAAACATCATCGGCGGCCAACATGTGACCCAGCTGGCTGCACCCTG  666
         ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  601  GGCCCCACCCCCGTGAACATCATCGGCGGCCAACATGTGACCCAGCTGGCTGCACCCTG  660

Qy  667  AACTTCCCATCAGCCCATGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC  726
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  AACTTCCCATCAGCCCATGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC  720

Qy  727  CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGC  786
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGC  780

Qy  787  GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACC  846
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACC  840

Qy  847  CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC  906
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC  900

Qy  907  GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCCACCCGCGCC  966
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCCACCCGCGCC  960

Qy  967  GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGGGACGCCTACTTCAGCGTG  1026
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGGGACGCCTACTTCAGCGTG  1020

Qy 1027  CCCCTGGACGAGGACTTCCGCAAGTACACACCGCCTTCACCCTGCCCAGGGCTGCAACAACGAG  1086
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  CCCCTGGACGAGGACTTCCGCAAGTACACACCGCCTTCACCCTGCCCAGGGCTGCAACAACGAG  1080

Qy 1087  ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC  1146
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC  1140

Qy 1147  ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC  1206
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC  1200

Qy 1207  GTGATCTACCAGGGCCCCCTGTACGTGGGCAGGGACCTGGAGATGGGCCAGACCGCGC  1266
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  GTGATCTACCAGGGCCCCCTGTACGTGGGCAGGGACCTGGAGATGGGCCAGACCGCGC  1260

Qy 1267  AAGATCAGGAGAGCTGCGCAAGCACCTGCTGCGCCTGGGCTTCACCACCCCGACAAGAAG  1326
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  AAGATCAGGAGAGCTGCGCAAGCACCTGCTGCGCCTGGGCTTCACCACCCCGACAAGAAG  1320

Qy 1327  CACCAGAAGAGCCCCTTCGTGTGATGGGCTACAGCGTCACCCGACAGAGTGACC  1386
         ||||||||||||||||||||||||||          |||||||||||||||||
Db 1321  CACCAGAAGAGCCCCTTCGTGTGCCCAT------CGAGCTGCACCCCGACAAGAGTGACC  1374

Qy 1387  GTGCAGCCCATGAGCTGCCCGAGAAGGAGAGACTGAACCGTGAACGTGAAGACATCCAGAGCTG  1446
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1375  GTGCAGCCCATGAGCTGCCCGAGAAGGAGAGACTGAACCGTGAACGTGAAGACATCCAGAGCTG  1434

Qy 1447  GTGGGCAAGCTGAACTGGGCGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAGCTGTGC  1506
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1435  GTGGGCAAGCTGAACTGGGCGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAGCTGTGC  1494

Qy 1507  AAGTGCTGCGGCGCCCAAGGCCCTGACCGACATGTGCCCCTGACCGAGGAGGCCGAG  1566
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1495  AAGTGCTGCGGCGCCCAAGGCCCTGACCGACATGTGCCCCTGACCGAGGAGGCCGAG  1554

Qy 1567  CTGGAGCTGGCCGAGAACCGGCCGAGATCCTGCGCGAGCCGTGCACGGCGTGTACTACGAC  1626
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1555  CTGGAGCTGGCCGAGAACCGGCCGAGATCCTGCGCGAGCCGTGCACGGCGTGTACTACGAC  1614

Qy 1627  CCCAGCAAGGACCCGGTGCCGGATCCAGATCCAGGAGCAGGGCCACGACCAGTGACCTACCAG  1686
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1615  CCCAGCAAGGACCCGGTGCCGGATCCAGATCCAGGAGCAGGGCCACGACCAGTGACCTACCAG  1674

Qy 1687  ATCTACCAGGAGCCCTTCAAGAACCTGAAGAACCTGAAGAGACCGGCAAGTACGCGCACCGCC  1746
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
;  OTHER INFORMATION: p2Pol.opt.YMWM_C
US-10-190-305A-37

Query Match             98.4%;  Score 2422.6;  DB 14;  Length 2445;
Best Local Similarity   99.6%;  Pred. No. 0;
Matches 2441;  Conservative   0;  Mismatches   4;  Indels   6;  Gaps   1;

Qy       7 GCCACCATGGCCGAGGCCATGAGCGAGGCCAGGCCAACATCCTGATGCAGGCCAGC     66
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GCCACCATGGCCGAGGCCATGAGCGAGGCCAGGCCAACATCCTGATGCAGGCCAGC     60

Qy      67 AACTTCAAGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCATC    126
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AACTTCAAGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC    120

Qy     127 GCCCGGAACTGCCGGCCCCCGCCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC    186
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCCCGGAACTGCCGGCCCCCGCCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC    180

Qy     187 CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCC    246
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCC    240

Qy     247 CAGGGCAAGGCCCGGGAGTTCCCCAGCGAGACCGAGAACCGCGGCCAACAGCCCACCAGCCGAC    306
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAGGGCAAGGCCCGGGAGTTCCCCAGCGAGACCGAGAACCGCGGCCAACAGCCCACCAGCCGAC    300

Qy     307 GAGCTGCAGGTGGCGGGACAACCCCCGGCGGAGGCGGAGGCGGGCCAGGGCACC    366
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GAGCTGCAGGTGGCGGGACAACCCCCGGCGGAGGCGGAGGCGGGCCAGGGCACC    360

Qy     367 CTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGTGGGCGGC    426
           |||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db     361 CTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGTGGGCGGC    420

Qy     427 CAGATCAAGGAGCCCTGCTGGCACACGGCCCGAGGACACCGTGTGAGGAGATGAGC    486
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CAGATCAAGGAGCCCTGCTGGCACACGGCCCGAGGACACCGTGTGAGGAGATGAGC    480

Qy     487 CTGCCCGGCAGAGTGGAGCCCAAGATGCCATCGCCGGCTTCATCAAGGTGCGC    546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTGCCCGGCAGAGTGGAGCCCAAGATGCCATCGCCGGCTTCATCAAGGTGCGC    540

Qy     547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC    606
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC    600

Qy     607 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG    666
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG    660

Qy     667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC    726
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC    720

Qy     727 CCCAAGGTGAAGCAGTGCCCCTGACCGAGGAGAATCAAGGCCCTGACCGACCCATCTGC    786
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCCAAGGTGAAGCAGTGCCCCTGACCGAGGAGAATCAAGGCCCTGACCGACCCATCTGC    780

Qy     787 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACACACC    846
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC    840

Qy     847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC    906
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC    900

Qy     907 GAGCTGAACAAGGCACCCAGGACTTCTGGGAGGTGCATCGGCATCCCCACCCCGCC    966
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAGCTGAACAAGGCACCCAGGACTTCTGGGAGGTGCATCGGCATCCCCACCCCGCC    960

Qy     967 GGGCTGAGGAAGAAGAGAGAGCGTGACCGTGCTGGACGTGGGCGAACGCCTACTTCAGCGTG    1026
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1675 ATCTACCAGGAGCCCCTTCAAGAACCTGAAGACCGGCCAAGTACGCCAAGATGCGCACCGCC    1734
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    1747 CACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCTGCAGAAGATCGCCAGGAGCATC    1806
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1735 CACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCTGCAGAAGATCGCCAGGAGCATC    1794

Qy    1807 GTGATCTGGGCAAGACCCCAAGTTCCGCCTGCCATCCAGAAGGAGACCTGGGAGACC    1866
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1795 GTGATCTGGGCAAGACCCCAAGTTCCGCCTGCCATCCAGAAGGAGACCTGGGAGACC    1854

Qy    1867 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC    1926
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1855 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC    1914

Qy    1927 CCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGGCGCCGAGACCTTC    1986
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1915 CCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGGCGCCGAGACCTTC    1974

Qy    1987 TACGTGGACGGCGCCCAACCGCGAGACAAGATCGGCAAGGCCGGCTACGTGACCGAC    2046
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1975 TACGTGGACGGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGAC    2034

Qy    2047 CGGGGGCGGCGAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG    2106
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2035 CGGGGGCGGCGAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG    2094

Qy    2107 GCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGGAGGTGAACATCGTGACCGACAGCCAG    2166
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2095 GCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGGAGGTGTAACATCGTGACCGACAGCCAG    2154

Qy    2167 TACGCCCTGGGCATCATCCAGGCCCCGACAAGAGACGGCGAGCTGGTGAACCAG    2226
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2155 TACGCCCTGGGCATCATCCAGGCCCCGACAAGAGACGGCGAGCTGGTGAACCAG    2214

Qy    2227 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACTTGAGCTGGGTGCCCGCCACAAG    2286
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2215 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACTTGAGCTGGGTGCCCGCCACAAG    2274

Qy    2287 GGCATCGGCGCAACGAGCAGATCGACAACGACGAGATCCGTGAGCAAGGGCATCGCCAAGGTGCTG    2346
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2275 GGCATCGGCGCAACGAGCAGATCGACAACGACGAGATCCGTGAGCAAGGGCATCGCCAAGGTGCTG    2334

Qy    2347 TTCCTGGACGGCATCGATGCGCGCCATCGTGATCTACCAGTACATGACGACCTGTACGTG    2406
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2335 TTCCTGGACGGCATCGATGCGCGCCATCGTGATCTACCAGTACATGACGACCTGTACGTG    2394

Qy    2407 GGCAGCGGCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT    2457
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2395 GGCAGCGGCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT    2445
```

RESULT 6
US-10-190-305A-37
; Sequence 37, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:

```
Db   961 GGCCTGAAGAAGAAGAGCGTGACCGTGTGGGCGACGGCCTACTTCAGCGTG 1020

Qy  1027 CCCCTGGACGACTTCCGCCAAGTACACCGCCTTCACCATCCAACAACGAG 1086
         |||||||||||||||||||||||||  |||||||||||||||||||||||
Db  1021 CCCCTGGACGACTTCCGCCAAGTACACCGCCTTCACCATCCAACAACGAG 1080

Qy  1087 ACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGC 1146
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGC 1140

Qy  1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCCCAGATC 1206
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCCCAGATC 1200

Qy  1207 GTGATCTACCAGGCCCCCTGTACGTGGGCAGCGCGACCTGGCAGCCAGC 1266
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTGATCTACCAGGCCCCCCTGTACGTGGGCAGCGCGACCTGGCAGCCAGC 1260

Qy  1267 AAGATCGAGGAGCAGCATGCCAAGCACCTGCTGCCGCCTTCACCCCCGACAGAAG 1326
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AAGATCGAGGAGCAGCATGCCAAGCACCTGCTGCCGCTGGGGCTTCACCACCCCGACAAGAAG 1320

Qy  1327 CACCAGGAAGGAGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGACC 1386
         |||||||||||||||||||||||||     |||||||||||||
Db  1321 CACCAGGAAGGAGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACAAGTGACC 1374

Qy  1387 GTGCAGCCCATGAGCTGCCCAGAAGGAGAGTGACCGTGAACGATCCAGAGCTG 1446
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1375 GTGCAGCCCATCGAGCTGCCCAGAAGGAGAGTGACCGTGAACGATCCAGAGCTG 1434

Qy  1447 GTGGCAAGCTGAACTGGCCAGCCAGATCTACCCCGCATCAAGGTGCGCCAGCTGTGC 1506
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1435 GTGGCAAGCTGAACTGGCCAGCCAGATCTACCCCGCATCAAGGTGCGCCAGCTGTGC 1494

Qy  1507 AAGGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGGCCGAG 1566
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1495 AAGGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGGCCGAG 1554

Qy  1567 CTGGAGTGGCCGAGAACCCGGAGATCCGCGCGAGCCCGTGCACGGCGTGCTGTACTACGAC 1626
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1555 CTGGAGTGGCCGAGAACCCGGAGATCCGCGCGAGCCCGTGCACGGCGTGCTGTACTACGAC 1614

Qy  1627 CCCAGCAAGGACCTGGTGCGCCGGATCCAGAAGCAGCAGGGCCACCAGTCGGACCTACCAG 1686
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1615 CCCAGCAAGGACCTGGTGCGCCGGATCCAGAAGCAGCAGGGCCACCAGTCGGACCTACCAG 1674

Qy  1687 ATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCC 1746
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1675 ATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCC 1734

Qy  1747 CACACCAACGACGTGAAGCAGTGAAGGAGAACCTGGAGAGATCGCCGTGCAGAAGATCGCC 1806
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1735 CACACCAACGACGTGAAGCAGTGAAGGAGAACCTGGAGAGATCGCCGTGCAGAAGATCGCC 1794

Qy  1807 GTGATCTGGGCCAAGACCCCAAGCCGCCATGGGACCTGGGAGACC 1866
         |||||||||||||||||||||||||||||
Db  1795 GTGATCTGGGCCAAGACCCCCAAGCCGCCATGGGACCTGGGAGACC 1854

Qy  1867 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC 1926
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1855 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCC 1914

Qy  1927 CCCCTGGTGAAGCTGTGTACCAGCTGACGGAAGAAGATCGGCGGCGAAGGTGCGGCACCAGATGA 1986
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1915 CCCCTGGTGAAGCTGTGTACCAGCTGACGGAAGAAGATCGGCGGCGAAGGTGCGGCACCAGATGA 1974

Qy  1987 TACGTGGACCGGCCGCGCCAACCAGCCGCGCTACGTGCACGAC 2046
         |||||||||||||||||||||||||||||
Db  1975 TACGTGGACCGGCCGCGCCAACCAGCCGCGCTACGTGCACGAC 2034

Qy  2047 CGGGGCCGGCAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG 2106
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2035 CGGGGCCGGCAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAG 2094
```

```
Qy  2107 GCCATCCAGCTGGCCCTGCGAGGACGCAGCGGCAGCGGCGAGGTGAACATGTGACCGACAGCGCAG 2166
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2095 GCCATCCAGCTGGCCCTGCGAGGACGCAGCGGCAGCGGCGAGGTGAACATCGTGACCGACAGCCCAG 2154

Qy  2167 TACGCCCTGGCATCATCCAGGCCCGACAAGAGCGAGCTGGTGAACCAG 2226
         |||||||||||||||||||||||||||||||||||||||||
Db  2155 TACGCCCTGGCATCATCCAGGCCCGACAAGAGCGAGCTGGTGAACCAG 2214

Qy  2227 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGACCTGGTGCCGCCCACAAG 2286
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2215 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGACCTGGTGCCGCCCACAAG 2274

Qy  2287 GGCATCGGCGCAACGAGCAAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTG 2346
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2275 GGCATCGGCGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTG 2334

Qy  2347 TTCCTGGACGGCATCGATGGCGCCATCGATCGTGATCGTGATCGACGACCCTGTACGTG 2406
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2335 TTCCTGGACGGCATCGATGGCGCCATCGATCGTGATCGTGATCGACCCTGTACGTG 2394

Qy  2407 GGCAGCGGCGGCGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCCGGT 2457
         |||||||||||||||||||||||||||||||||||||||
Db  2395 GGCAGCGGCGGCGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCCGGT 2445
```

RESULT 7
US-10-190-435-9

; Sequence 9, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmut_C
US-10-190-435-9

Query Match   98.1%;   Score 2415.6;   DB 14;   Length 3930;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 2434; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

```
Qy    14 TGGCCAGGGCCATGAGCCAGGCCACCAGCGCCACAATCCTGATGCGCAGCAGCAACTTCA 73
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1487 TCCCCGAGGCCATGAGCCAGGCCACCAGCGCCACAATCCTGATGCGCAGCAGCAACTTCA 1546

Qy    74 AGGGCCCCAAGGCGCATCATCAAGTGGTTCAACTCGCGCAAGAGGGCCACATCGCCCCGCA 133
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1547 AGGGCCCCAAGGCGCATCATCAAGTGGTTCAACTCGCGCAAGAGGGCCACATCGCCCCGCA 1606

Qy   134 ACTGCGCGCCGACCGGCCGCAACTTCTTCCGGCAAGGTGCGGCAAGGAGGGCAAGGAGGG 193
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1607 ACTGCGCGCCGACCGGCCGCAACTTCTTCCGGCAAGGTGCGGCAAGGAGGGCAAGGAGGG 1666

Qy   194 AGGACTGCACCGAGCGCCAGGCCAACCGCCGCCAACTTCTTCCGGGGCCTTCCCCAGGGCA 253
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1667 AGGACTGCACCGAGCGCCAGGCCAACCGCCGCCAACTTCTTCCGGGGCCTTCCCCAGGGCA 1726

Qy   254 AGGCCCGGCAGTTCCCCAGCGCCGAGAGACCGCGCCAACAGCCGCGAGCTGC 313
         |||||||||||||||||||||||||||||||||
```

```
Db 1727 AGGCCCGCGGAGTTCCCAGCGACGAGACCGCGCCAACAGCCCACCAGCCGCGAGCTGC 1786
         ||||||||||||||||||||||||||||||||||  |||||||||||||||||||||
Qy  314 AGGTGCGCGGCGACAACCCCGCAGGGAGGCGCCGGCGCGGAGGCGCCAGGGCACCCTGAACT 373

Db 1787 AGGTGCGCGGCGACAACCCCGCAGGGAGGCGCCGGCGCGGAGGCGCCAGGGCACCCTGAACT 1846
Qy  374 TCCCCCAGATCACCCGTGTGGCAGCGCCCCCTGGTGAGCGATCAAGGTGGGCGGCCAGATCA 433

Db 1847 TCCCCCAGATCACCCGTGTGGCAGCGCCCCCTGGTGAGCGATCAAGGTGGGCGGCCAGATCA 1906
Qy  434 AGGAGGCCCTGCTGGACACCGCCGCGACGACACCGGCACCGGTGTACTAGACCCCAGCA 493

Db 1907 AGGAGGCCCTGCTGGACACCGCCGCGACGACGACCGGCGTGGAGGATGAGCTGCCCG 1966
Qy  494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 553

Db 1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 2026
Qy  554 ACCAGATCCTGATGAGATCGCGCGCAAGAGGCCATCGGCCAAGAGGGCCATCGGCCTGATCGGCCCCA 613

Db 2027 ACCAGATCCTGATGAGATCGCGCGCAAGAGGCCATCGGCCAAGAGGGCCATCGGCCTGATCGGCCCCA 2086
Qy  614 CCCCCGTGAACATCATCGGCCGCGCAACATGCTGACCCCAGCTGGGCTGCACCCTGAACTTCC 673

Db 2087 CCCCCGTGAACATCATCGGCCGCGCAACATGCTGACCCCAGCTGGGCTGCACCCTGAACTTCC 2146
Qy  674 CCATCAGCCCCATCCAGACCCCTGACCGGCGAAGCTGAAGCCCGGACAGGGCCCCCAAGG 733

Db 2147 CCATCAGCCCCATCCAGACCCCTGACCGGCGAAGCTGAAGCCCGGACAGGGCCCCAAGG 2206
Qy  734 TGAAGCAGTGCCCCTGACCGGAGAAGACAGCACCAGCAGCAGCAACAGGCCCCGAGTTCGCGCCCCGTGT 793

Db 2207 TGGAGAAGGAGGGCAAGACACCACCAAGAGCACCAGCACCAAGAGGCCCCGAGACCCCTACACACCCCGTGT 2266
Qy  794 TGGGAAGGAGGGCAAGACACCACCAAGAGCACCAGCACCAAGAGGCCCCGAGACCCTACACACCCCGTGT 853

Db 2267 TGGGAGGAGGGCAAGACACCACCAAGAGCACCAGCACCAAGAGGCCCCGAGACCCTACACACCCCGTGT 2326
Qy  854 TCGGCCATCAAGAGGACGACGCACCCAAGTGGCGCGCAAGTGCGCGCAAGTGCGCGAGCTTCCGCGAGCTGA 913

Db 2327 TCGCCATCAAGAGGACGACGCACCCAAGTGGCGCGCAAGTGCGCGCAAGTGGCGAGCTTCCGCGAGCTGA 2386
Qy  914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCGCCCGGCCGCCGGCCTGA 973

Db 2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCGCCCGCCGGCCTGA 2446
Qy  974 AGAAGAAGAAGGAGCGGTGACCGGCTGACGTGCCGCTTGACGTGGGGCGACGCCACGCGGCCATCC 1033

Db 2447 AGAAGAAGAAGGAGCGGTGACCGGCTGACGTGCCGCTTGACGTGGGGCGACGCCACGCGGCCATCC 2506
Qy 1034 ACGAGGACTTCCGCAAGCATCCCCGCCCGGCATCCCCAGCATCAACAACAACAGAGACCCCCGG 1093

Db 2507 ACGAGGACTTCCGCAAGCATCCCCGCCCGGCTTCACCATCCCAGCATCAACAACAACAGAGACCCCCG 2566
Qy 1094 GCATCCGCTACCAGGTGCTGCCCCGAGCTGCCCGCGCCTCCGGCCGCCTTCCTCCCGCGCCAGATCTT 1153

Db 2567 GCATCCGCTACCAGGTGCTGCCCCGAGCTGCCCGCGCCTCCGGCCGCCTTCCTCCCGCGCCAGATCTT 2626
Qy 1154 AGAGGACGCATGACCAAGATCCGAGATCGTGATCT 1213

Db 2627 AGAGGACGCATGACCAAGATCCGAGATCGTGATCT 2686
Qy 1214 ACCAGGGCCCCCCTGTACGTGGGCAGCTGACCTTCACCCACCAGCACCGCGCAAGATCG 1273

Db 2687 ACCAGGGCCCCCTGTACGTGGGCAGCTGACCTTCACCCACCAGCCGCGCAAGATCG 2746
Qy 1274 AGGAGCTGCGCAAGCACCTGCTGCGCTTGCGCGTCTGCGCGTGTGGGGCTTCACCACGAGAAGCACCAGA 1333

Db 2747 AGGAGCTGCGCAAGCACCTGCTGCGCTTGCGCGTCTGCGCGTGTGGGGCTTCACCACGAGAAGCACCAGA 2806
Qy 1334 AGGAGCCCCCTTCCTTGTGGATGGCGCCTACGAGCTGCACGCCCAGTGACCGTGCAGC 1393

Db 2807 AGGAGCCCCCTTCCTTGTGGATGGCGCCTACGAGCTGCACGCCCAGTGACCGTGCAGC 2860
```

```
Db 1394 CCATCGAGCTGCCCGAGAAGGAGGACTGGACCGTGAACGACATCCAGGAGCTGGTGGGCA 1453
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 2861 CCATCGAGCTGCCCGAGAAGGAGGACTGGACCGTGAACGACATCCAGGAGCTGGTGGGCA 2920

Db 1454 AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCGGCACCGAGCTGTCAGCTGC 1513
Qy 2921 AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCGGCACCGAGCTGTGCAAGCTGC 2980

Db 1514 TGCGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCCAGCTGGGAGC 1573
Qy 2981 TGCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCGAGCTGGGAGC 3040

Db 1574 TGGCCGAGAACCGCAGATCTGCGGCAGCCGTGCACGGCGTGTACTAGACCCCAGCA 1633
Qy 3041 TGGCCGAGAACCGCGAGATCCTGCGGCAGCCGTGCACGGCGTGTACTACGACCCCAGCA 3100

Db 1634 AGGACTTGGTGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACTACCAGATCTACC 1693
Qy 3101 AGGACTTGGTGCCCAGATCCAGAAGCAGGGCCACGACCAGTGGACTTACCAGATCTACC 3160

Db 1694 AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCAAGACCGCCACACCA 1753
Qy 3161 AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCAAGACCGCCACACCA 3220

Db 1754 ACGACGTGAAGCAGTCCGCTGACGAGGCCGTGCCGACTCGCCATCGAGAGATCGTGTGATCT 1813
Qy 3221 ACGACGTGAAGCAGTCCGCTGACGAGGCCGTGCCGACTCGCCATCGAGAGATCGTGTGATCT 3280

Db 1814 GGGGCAAGACCCCCAAGTTCCGCGCTGATCCCGAGGTGGAGACCTGGTGGA 1873
Qy 3281 GGGGCAAGACCCCCAAGTTCCGCGCTGATCCCGAGGTGGAGACCTGGTGGA 3340

Db 1874 CCGACTACTGGCGAGGCCACCTGATCCCCGAGGTGGGAGTTCGTGAACACCCCCTGG 1933
Qy 3341 CCGACTACTGGCGAGGCCACCTGATCCCCGAGGTGGGAGTTCGTGAACACCCCCTGG 3400

Db 1934 TGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCCAGCCTTCTACGTGG 1993
Qy 3401 TGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCCAGCCTTCTACGTGG 3460

Db 1994 ACGGCGCCGCCAACCGCCCAGATCGGCGAGACCAAGATCGGCCGGCTACGTGACCGGCGGGGCC 2053
Qy 3461 ACGGCGCCGCCAACCGCCCAGATCGGCGAGACCAAGATCGGCCGGCTACGTGACCGGCGGGGCC 3520

Db 2054 GGCAGAAGATCGTGAGCCTGACCGGAGGTGAACATCGTGACCGAGCTGCAGGCCATCC 2113
Qy 3521 GGCAGAAGATCGTGAGCCTGACCGGAGGTGAACATCGTGACCGAGCTGCAGGCCATCC 3580

Db 2114 AGCTTGGCCCTGCAGGAGCAGCGGCCAGCAGCGCAACATCGTGACCGAGCCAGTACGCCC 2173
Qy 3581 AGCTTGGCCCTGCAGGAGCAGCGGCCAGCAGCGCAACATCGTGACCGAGCCAGTACGCCC 3640

Db 2174 TGGGCATCATCCAGGCCCGACAGAGGAGAGCCCAGCGAGCTGGTGGAACCAGATCATCG 2233
Qy 3641 TGGGCATCATCCAGGCCCGACAGAGGAGAGCCCAGCGAGCTGGTGGAACCAGATCATCG 3700

Db 2234 AGCAGCTGATCGATCAAGAGCAGCAGATCGACAACAGCTGGTGACAAGCTGGTGAGCGGGCATCG 2293
Qy 3701 AGCAGCTGATCGATCAAGAGCAGCAGATCGACAACAGCTGGTGACAAGCTGGTGAGCGGGCATCG 3760

Db 2294 GCGGCAACGACAGCGCATCGATGGCGCATCGTGATCGCCAGTACGAGCTGTGTACGGCGACGCG 2353
Qy 3761 GCGGCAACGACAGCGCATCGATGGCGCATCGTGATCGCCAGTACGAGCTGTGTACGGCGACGCG 3820

Db 2354 ACGGCATCGATGGCGCATCGATGGCGCATCGTGATCGCCAGTACGAGCTGTGTACGGCGACGCG 2413
Qy 3821 ACGGCATCGATGGCGCATCGATGGCGCATCGTGATCGCCAGTACGAGCTGTGTACGGCGCAGCG 3880

Db 2414 GCGGGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
Qy 3881 GCGGGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 3924
```

RESULT 8
US-10-190-435-10
; Sequence 10, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: Patentin Ver. 2.0
;
; SEQ ID NO 10
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutAtt_C

US-10-190-435-10

Query Match   98.0%; Score 2414; DB 14; Length 3930;
Best Local Similarity   99.5%; Pred. No. 0;
Matches 2433; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA   73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 TCGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA 1546

Qy   74 AGGGCCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCA 133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 AGGGCCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCA 1606

Qy   134 ACTGCGGCGCCCCCCGCAAGAAGGGCGTGCTGGAAGTCGGCGCAAGAAGGGCCACCAGATGA 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 ACTGCGGCGCCCCCCGCAAGAAGGGCGTGCTGGAAGTCGGCGCAAGAAGGGCCACCAGATGA 1666

Qy   194 AGGACTGCACCGAGCGCGCCAGGCCAACTTCTTCGCGGAGGACCTGGCCTTCCCCCAGGGCA 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667 AGGACTGCACCGAGCGCGCCAGGCCAACTTCTTCGCGGAGGACCTGGCCTTCCCCCAGGGCA 1726

Qy   254 AGGCCCGCGAGTTCCCCAGCAGCGAGCAGACACCGGCCAACAGCGAGCCGCGAGCTGC 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727 AGGCCCGCGAGTTCCCCAGCAGCGAGCAGACACCGGCCAACAGCCCGCCAGCCGCGAGCTGC 1786

Qy   314 AGTGCGCGGCGGACAACCCCCGAGCGGCCGGCCGCGGACGCCGAGCCGCCAGGCCACCTGAACT 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787 AGTGCGCGGCGGACAACCCCCGAGCGGCCGGCCGCGGACGCCGAGCCGCCAGGCCACCTGAACT 1846

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGTGAGTCGGCAGCATCATCAAGGTGTCGGCGGCCAGATCA 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847 TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGTGAGTCGGCAGCATCATCAAGGTGTCGGCGGCCAGATCA 1906

Qy   434 AGGAGGCCCTGCTTCGGGAACGGCCGACGCACGACCACCGTGCTGCTGTGCTGAGCACCGGTACTACGGAGACCCCAGCA 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907 AGGAGGCCCTGCTTCGGGAACGGCCGACGCACGACCACCGTGCTGCTGTGCTGAGCACCGGTACTACGGAGACCCCCAGCA 1966

Qy   494 GCAAGTGGAAGCCAAGATGATCGGCGGCCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967 GCAAGTGGAAGCCAAGATGATCGGCGGCCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 2026

Qy   554 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGCCACCGTGCTGTCGGCCCCCA 613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGCCACCGTGCTGTCGGCCCCCA 2086

Qy   614 CCCCCGTGAACATCATCGGCCGCAACATCGCCGCCGCAACATGCTGACCCCAGTCGACCCCTGAACTTCC 673
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2087 CCCCCGTGAACATCATCGGCCGCAACATCGCCGCCGCAACATGCTGACCCCAGTCGACCCCTGAACTTCC 2146

Qy   674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 733
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2147 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 2206

Qy   734 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 793
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2207 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 2266

Qy   794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACACACACCCCGTGT 853
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2267 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACACACACCCCGTGT 2326

Qy   854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 2386

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCCACCCCGCCGGCCTGA 973
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCCACCCCGCCGGCCTGA 2446

Qy   974 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2447 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 2506

Qy   1034 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 1093
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2507 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 2566

Qy   1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCTGGAGGGCAGCATCTTCC 1153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2567 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCTGGAGGGCAGCATCTTCC 2626

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAATGATCGTGATCT 1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAATGATCGTGATCT 2686

Qy   1214 ACCAGCGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGGCCCAAGATCG 1273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2687 ACCAGCGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGGCCCAAGATCG 2746

Qy   1274 AGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 1333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2747 AGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 2806

Qy   1334 AGGAGCCCCCTTCCTTCGTGTGATGGGCTACGAGTTGCACCCCGACAAGTGCACCGTGCAGC 1393
           |||||||||||||||||||||||
Db   2807 AGGAGCCCCCTTCCTTCGTGTGCCCAT------CGAGCTGCACCCCGACAAGTGCACCGTGCAGC 2860

Qy   1394 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCA 1453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCA 2920

Qy   1454 AGTGAACTGGGCCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC 1513
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGTGAACTGGGCCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGC 2980

Qy   1514 TCCCCCAGATCACCCTGTGGCAGCACCGGGCCGCCGCCCTGACCGAGATCGTGCCCCTGAGAGC 1573
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 TGCGCGGCGCAAGGCCGCGAGATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGC 3040

Qy   1574 TGGCCGAGAACCGGCAGATCGTGCCCCTGACCGAGATCGTGCCCCTGACGCGGCGTGTACTAGACCCCAGCA 1633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 TGGCCGAGAACCGGCAGATCGTGCCCCTGACCGAGATCGTGCCCCTGACGCGGCGTGTACTACGACCCCAGCA 3100

Qy   1634 AGGAACTTGGTGGCGGGATCTGCGCCAAGAGGCAGCAGGCCACGAGCAGTGGACCTACC 1693
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 AGGACTTGGTGGCGGGATCTGCGCCAAGAGGCAGCAGGCCACCAGTGGACCTACC 3160

Qy   1694 AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATGGCACCGCCCACACCA 1753
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATGGCACCGCCCACACCA 3220

Query Match   98.0%;   Score 2414;   DB 14;   Length 3930;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2433; Mismatches 0; Conservative 0; Indels 5; Gaps 1;

```
Qy    14 TGGCCGAGGCATGAGCCAGGCAGGCCAACAATCCTGATGCAGCGCGAGCAACTTCA   73
         |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db  1487 TCGCCGAGGCATGAGCCAGGCAGGCCAACAATCCTGATGCAGCGCGAGCAACTTCA 1546

Qy    74 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCAGATGA  133
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1547 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCAGATGA 1606

Qy   134 ACTGCCGGCGCCCCGGCCAAGAAGGGTGCTGAAGTGCGGCAAGGAGGGCCAGGGGCA  193
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1607 ACTGCCGGCGCCCCGGCCAAGAAGGGTGCTGAAGTGCGGCAAGGAGGGCCAGGGGCA 1666

Qy   194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA  253
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1667 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA 1726

Qy   254 AGGCCCCGCGAGTTCCCCAGCAGCAGAACCGCGCCAACAGCCCCACCAGCCCGCGAGCTGC  313
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1727 AGGCCCCGCGAGTTCCCCAGCAGCAGAACCGCGCCAACAGCCCCACCAGCCCGCGAGCTGC 1786

Qy   314 AGTGCGCGGCGACAACCCCCCAGCAGCCCGCGAGCGCCACCCTGAACT  373
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1787 AGTGCGCGGCGACAACCCCCCAGCAGCCCGCGAGCGCCACCCTGAACT 1846

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGCCCCGTGGCATCAAGGTGGGCCCGCCCAGATCA  433
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1847 TCCCCCAGATCACCCTGTGGCAGCGCCCCGTGGCATCAAGGTGGGCCCGCCCAGATCA 1906

Qy   434 AGGAGGCCCTGGCACGCGACACCGACACCGCATCGGCGGCAAGAAGGCCATCGGCC  493
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1907 AGGAGGCCCTGGCACGCGACACCGACACCGCATCGGCGGCAAGAAGGCCATCGGCC 1966

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGGTGCGGCCCAGTACG  553
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGGTGCGGCCCAGTACG 2026

Qy   554 ACCAGATCCTGATCGAGATCTGCGCCGCAACATGCTGACCCAGTCGGCCCCA  613
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2027 ACCAGATCCTGATCGAGATCTGCGCCGCAACATGCTGACCCAGTCGGCCCCA 2086

Qy   614 CCCCCGTGAACATCATCGGGGCTGCCCAGTCGGGCTGCACCCTGAACTTCC  673
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2087 CCCCCGTGAACATCATCGGGGCTGCCCAGTCGGGCTGCACCCTGAACTTCC 2146

Qy   674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCAGGCCCCAAGG  733
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2147 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCAGGCCCCAAGG 2206

Qy   734 TGAAGCAGTGCCCCTGACCGGAGAGATCAAGGCCCTGACCGGCATCGGAGAGAGA  793
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2207 TGAAGCAGTGCCCCTGACCGGAGAGATCAAGGCCCTGACCGGCATCGGAGAGAGA 2266

Qy   794 TGGAAGGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCGTGT  853
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  2267 TGGAAGGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCGTGT 2326

Qy   854 TCGCCATCAAGAAGGACACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  913
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2327 TCGCCCATCAAGAAGGACACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 2386

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCTGA  973
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCTGA 2446

Qy   974 AGAAGAAGAAGAGAGTGTGAGACTGGCCGACGTGCTGACACTTCAGCGTGCCCCTG 1033
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2447 AGAAGAAGAAGAGAGTGTGAGACTGGCCGACGTGCTGACACTTCAGCGTGCCCCTG 2506
```

```
Qy  1754 ACGACGTGAAGCAGCTGACCGAGGCCGTCGTGCGAGAGATCGCCATCGAGAGCATCGTGATCT 1813
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3221 ACGACGTGAAGCAGCTGACCGAGGCCGTCGTGCGAGAGATCGCCATCGAGAGCATCGTGATCT 3280

Qy  1814 GGGCCAAGACCCCCAAGTTCCGCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGA 1873
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3281 GGGCCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGA 3340

Qy  1874 CCGACTACTGGCAGGCCACCTGACATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGG 1933
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3341 CCGACTACTGGCAGGCCACCTGACATCCCCGAGTGGGAGTTCGTGAACACCCCCCCTGG 3400

Qy  1934 TGAAGCTGTGTACCAGTGAGAAGGAGCCCATCATCCGCGCCGAGACTTCTACGTGG 1993
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3401 TGAAGCTGTGTACCAGTGAGAAGGAGCCCATCATCCGCGCCGAGACTTCTACGTGG 3460

Qy  1994 ACGGCGCCGCCAACCGAGACCGAGATCGGCCAAGCCGGCTACGTGACCGACCGGGGCC 2053
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3461 ACGGCGCCGCCAACCGAGACCGAGATCGGCCAAGCCGGCTACGTGACCGACCGGGGGC 3520

Qy  2054 GGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCC 2113
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3521 GGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCC 3580

Qy  2114 AGCTGGCCCTGCAGGACAGCGGCGAGCGAGCAAGATCGTGACGACCAGCAGCCAGTACGCCC 2173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3581 AGCTGGCCCTGCAGGACAGCGGCGAGCGAGCAAGATCGTGACGACCAGCAGCCAGTACGCCC 3640

Qy  2174 TGGGGCATCATCCAGGCCCAGCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCG 2233
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3641 TGGGGCATCATCCAGGCCCAGCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCG 3700

Qy  2234 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG 2293
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3701 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG 3760

Qy  2294 GCGGCAACGAGCAGATCGACAAGCTTGGTGATGGCGGCCATCGCGCCAAGGTGCTGTTCCTGG 2353
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3761 GCGGCAACGAGCAGATCGACAAGCTTGGTGATGGCGGCCATCGCGCCAAGGTGCTGTTCCTGG 3820

Qy  2354 ACGGCATCGATGGCGGCATCACCAGTACATGGACGACTACTTGTACGTGGGGCAGCG 2413
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3821 ACGGCATCGATGGCGGCATCACCAGTACATGGACGACTACTTGTACGTGGGGCAGCG 3880

Qy  2414 GCGCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
         ||||||||||||||||||||||||||||||||||||||||||
Db  3881 GCGCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 3924
```

RESULT 9
US-10-190-435-11
; Sequence 11, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutIna_C
US-10-190-435-11

RESULT 10
US-10-190-435-58
; Sequence 58, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 5184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TatRevNefgagCpolina C
US-10-190-435-58

Query Match   98.0%; Score 2414; DB 14; Length 5184;
Best Local Similarity   99.5%; Pred. No. 0;
Matches 2433; Conservative   0; Mismatches   5; Indels   6; Gaps   1;

```
Db   4115 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAGAGCTGGTGGGCA 4174
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1454 AGCTGAACTGGGCCAGCCAGCCGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCGAACTGTCGCAAGCTGC 1513
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4175 AGCTGAACTGGGCCAGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCGAACTGTCGTGCAAGCTGC 4234

Qy   1514 TGCGCGGCGCCAAGGCCCTGACCATCGTGCCCCTGACCGCCGAGTGGGAGC 1573
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4235 TGCGCGGCGCCAAGGGCCCTGACCATCGTGCCCCTGACCGCCGAGTGGGAGC 4294

Qy   1574 TGGCCGAGAACCGGCGAGATCCTCGCGCAGCCCGTGCGCACGGGTGTACTAGACCCCAGCA 1633
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4295 TGGCCGAGAACCGGCGAGATCCTCGCGCAGCCCGTGCGCACGGGTGTACTAGACCCCAGCA 4354

Qy   1634 AGGACCTGGTGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCAGTGGA 1693
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4355 AGGACCTGGTGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCAGATCTACC 4414

Qy   1694 AGGAGCCCTTCAAGAACCTGACCGAGGCCCGTGCGAGAGATCGCCATCGTGAGAGCATCGTGATCT 1753
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4415 AGGAGCCCTTCAAGAACCTGACCGAGGCCCGTGCGAGAGATCGCCATCGTGAGAGCATCGTGATCT 4474

Qy   1754 ACGACGTGAAGCAGCGTGACCGAGGCCCATGGGGCAGGAGAGACCTGGGATGTGGA 1813
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4475 ACGACGTGAAGCAGCGTGACCGAGGCCCATGGGGCAGGAGAGACCTGGGATGTGGA 4534

Qy   1814 GGGGCAAGACCCCCAGTTCCGCCTGCCATCCAGAAGGAGACCTGGGATCGTGACCTGGTGGA 1873
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4535 GGGGCAAGACCCCCAGTTCCGCCTGCCATCCAGAAGGAGACCTGGGATCGTGACCTGGTGGA 4594

Qy   1874 CCGACTACTGGAGGCCACCTGGATCCCGTGGTCCCGGAGTTCGTGAACACCCCCCCTGG 1933
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4595 CCGACTACTGGAGGCCACCTGGATCCCGTGGTCCCGGAGTTCGTGAACACCCCCCCTGG 4654

Qy   1934 TGAAGCTGTGTTACCAGCTGGAGAAGAGAGCCCATCATCGGGCGCCAGCCAGTACGGTGG 1993
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4655 TGAAGCTGTGTTACCAGCTGGAGAAGAGAGCCCATCATCGGGCGCCAGCCAGTACGGTGG 4714

Qy   1994 ACGGCGCGCCAACCGCGAGATCGGCAGAGGCCAAGATCGGCGCCGGCTACGTGACCGACCGGGCC 2053
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4715 ACGGCGCGCCAACCGCGAGATCGGCAGAGGCCAAGATCGGCGCCGGCTACGTGACCGACCGGGGCC 4774

Qy   2054 GGCAGAAGATCGTGAGCCTGAGCGCAGCAGCGAGGTGAACATCGTGACCGAGCTGCAGGCCATCC 2113
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4775 GGCAGAAGATCGTGAGCCTGAGCGCAGCAGCGAGGTGAACATCGTGACCGAGCTGCAGGCCATCC 4834

Qy   2114 AGTTGGCCCTGCCAGGACAGCGCCAGCGACGAGGTGAACATCGTGACCGACCAGTACGCCC 2173
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4835 AGTTGGCCCTGCCAGGACAGCGCCAGCGACGAGGTGAACATCGTGACCGACCAGTACGCCC 4894

Qy   2174 TGGGCATCATCCAGGGCCAGGGCTTCGGAAGGCAGCGCCCAGAGGTGGTGAACCAGATCATCG 2233
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4895 TGGGCATCATCCAGGGCCAGGGCTTCGGAAGGCAGCGCCCAGAGGTGGTGAACCAGATCATCG 4954

Qy   2234 AGCAGCTGATCAAGAGAGAGGAGGTGTACCTGACCGTGGTGCCCGCCCCACAAGGGCATCG 2293
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4955 AGCAGCTGATCAAGAGAGAGGAGGTGTACCTGACCGTGGTGCCCGCCCCACAAGGGCATCG 5014

Qy   2294 GCGGCAACGACAGACGATCGGAACAAGCTGGTGAGCAAGCTGCGCAAGGTGCTGTTCCTGG 2353
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5015 GCGGCAACGACAGACGATCGGAACAAGCTGGTGAGCAAGCTGCGCAAGGTGCTGTTCCTGG 5074

Qy   2354 ACGGGCATCGATGGCGGCATCGTGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGG 2413
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5075 ACGGGCATCGATGGCGGCATCGTGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGGCAGCGG 5134

Qy   2414 GCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGGCTAGCACCGGT 2457
                ||||||||||||||||||||||||||||||||||||||||||||||
Db   5135 GCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGGCTAGCACCGGT 5178

RESULT 11
```

```
Qy    314 AGTGTGCGCGGCGACAACCCCGAGCGAGGCCGCGGCGAGGCCCAGGGCACCCTGAACT 373
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGTGTGCGCGGCGACAACCCCGAGCGAGGCCGCGGCGAGGCCCAGGGCACCCTGAACT 3100

Qy    374 TCCCCAGATCACCCTGTGCAGCGCCCCTGGTCAGCGCATCAAGGTGGGCGGCCAGATCA 433
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCAGATCACCCTGTGCAGCGCCCCTGGTGAGCGCATCAAGGTGGGCGGCCAGATCA 3160

Qy    434 AGGAGGGCCCTGCTGACACCGGCGCCGACAACCGGTGCTGAGGAGATGAGCCTGCCCG 493
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGGCCCTGCTGACACCGGCGCCGACAACCGGTGCTGAGGAGATGAGCCTGCCCG 3220

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG 553
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG 3280

Qy    554 ACCAGATCCTGATCGGAGATCTGCGCCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCA 613
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATCGGAGATCTGCGCCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCA 3340

Qy    614 CCCCGTGAACATCATCGGCCGTGCCCGTGAAGCTGGGTGCCACCCTGAACTTCC 673
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCGTGAACATCATCGGCCGTGCCCGTGAAGCTGGGTGCCACCCTGAACTTCC 3400

Qy    674 CCATCAGCCCCATCGAACCGCTGCCCGTGAAGCTGAAGCCGGGCATGGACGGGCCCCAAGG 733
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAACCGCTGCCCGTGAAGCTGAAGCCGGGCATGGACGGGCCCCAAGG 3460

Qy    734 TGAAGCAGTGGCCCCTGACCGAGAGAATCAAGGCCCTGACCGGCCATCTGCGAGAGA 793
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGGCCCCTGACCGAGAGAATCAAGGCCCTGACCGGCCATCTGCGAGAGA 3520

Qy    794 TGGAGGAGGGGCAAGATCACCACCAGAGACCCCTACACACCCCCCGTGT 853
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGAGGAGGGGCAAGATCACCACCAGAGACCCCTACACACCCCCCGTGT 3580

Qy    854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGGTGGCGCAAGCTTCCGGCAGCTGA 913
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGGTGGCGCAAGCTTCCGGCAGCTGA 3640

Qy    914 ACAAGCGCACCACCAGACTTCTGGGGAGTGCAGCTGCAGCTCCCCACCCGCCGCGGCCTGA 973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGCGCACCACCAGACTTCTGGGGAGTGCAGCTGCAGCTCCCCACCCGCCGCGGCCTGA 3700

Qy    974 AGAAGAAGAGAGCGGTGACCGTGACCGTGCTGCGTGACGGCGGGCCCCATCTCAGCGTGCCCCTGG 1033
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAAGAAGAGAGCGGTGACCGTGACCGTGCTGCGTGACGGCGGGCCCCATCTCAGCGTGCCCCTGG 3760

Qy   1034 ACGAGGACTTCGCAAGTACACCGGCCTTCACCATCCCCAGCATCAACAACGACGAGACCCCCCC 1093
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACGAGGACTTCGCAAGTACACCGGCCTTCACCATCCCCAGCATCAACAACGACGAGACCCCCCC 3820

Qy   1094 GCATCCGTACCAGTACACGTGCTCTGCCGACCCTGGAGGCCTGGCGGCGTGGAAGGGCAGCCCCAGCATCTTCC 1153
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGTACCAGTACACGTGCTCTGCCGACCCTGGAGGCCTGGCGGCGTGGAAGGGCAGCCCCAGCATCTTCC 3880

Qy   1154 AGGAGCAGCATGACCAAGGAATCCTGGAGCGCCCTTCCGGCCCCGCCAACCCGAGATCGTGATCT 1213
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3881 AGGAGCAGCATGACCAAGGAATCCTGGAGCGCCCTTCCGGCCCCGCCAACCCGAGATCGTGATCT 3940

Qy   1214 ACCAGGCCCCCCTGTACGTGGCAGCGGACCTGGCAGGCCTGGAGATCGGCGCCCAGGCCAAGATCG 1273
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3941 ACCAGGCCCCCCTGTACGTGGCAGCGGACCTGGCAGGCCTGGAGATCGGCGCCCAGGCCAAGATCG 4000

Qy   1274 AGGAGTCGCAAGCTGCGCAAGCACCTGTTCACCACCCCCGACAAGAGCACCACCAGA 1333
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4001 AGGAGTCGCAAGCTGCGCAAGCACCTGTTCACCACCCCCGACAAGAGCACCACCAGA 4060

Qy   1334 AGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTCACCGACCAAGTGGACCGTGCAGC 1393
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4061 AGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTCACCGACCAAGTGGACCGTGCAGC 4114

Qy   1394 CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAGAGCTGGTGGGCA 1453
```

```
US-10-190-305A-82
; Sequence 82, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 82
; LENGTH: 5184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: TatRevNefgagCpolina C
US-10-190-305A-82

Query Match             98.0%;  Score 2414;  DB 14;  Length 5184;
Best Local Similarity   99.5%;  Pred. No. 0;
Matches 2433;  Conservative  0;  Mismatches  0;  Indels  5;  Gaps  1;
```

```
Qy     14 TGGCGGAGGCCATGAGCCAGCCAGGCCACCAGCGCCACAGCATCCTGATGCAGCCAGCAACTTCA   73
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2741 TCGCGGAGGCCATGAGCCAGCCAGGCCACCAGCGCCACAGCATCCTGATGCAGCGCAGCAACTTCA 2800

Qy     74 AGGGCCCCCAAGGCCATCATCAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCGCA  133
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AGGGCCCCCAAGGCCATCATCAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCGCA 2860

Qy    134 ACTGCGGCCGCCCCCGCCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA  193
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 ACTGCGGCCGCCCCCGCCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAGGAGGGCCACCAGATGA 2920

Qy    194 AGGACTGACCAGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGCA  253
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGACTGACCAGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 2980

Qy    254 AGGGCCCGCGAGTTCCCCGCCAGCGACGAGAACCGGCGCCGCCGAGCGCCCAACAGCCGCGGAGCTGC  313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 AGCGCCCGCGAGTTCCCCGCCAGCGACGAGAACCGGCGCCGCCGAGCGCCCAACAGCCGCGGAGCTGC 3040

Qy    314 AGGTGCGCGGCGGACAACCCCCGCCAGCGCCCCGCCAGCGAGGCGCCGCCGAGCGCGCACACCCCTGTGACT  373
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGCGGCGGACAACCCCCGCCAGCGCCCCGCCAGCGAGGCGCCGCCGAGCGCGCACACCCCTGTGACT 3100

Qy    374 TCCCCAGATCACCCTGCTGGCAGGCGCCCCCTGTGAGCATCAAGGTGCGGCGGCGCCAGATCA  433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCAGATCACCCTGCTGGCAGGCGCCCCCTGTGAGCATCAAGGTGCGGCGGCGCCAGATCA 3160

Qy    434 AGGAGGCCCTGCTGGACACCGGCGCCGACGACAACCGGCGCCAGATGAGCCTGCCCG  493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGCCCTGCTGGACACCGGCGCCGACGACAACCGGCGCCAGATGAGCCTGCCCG 3220

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG  553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 3280

Qy    554 ACCAGATCCTGATCGAGATCTGCGGCCATAAGGCCATCGGCACCGTGCTGGTCGGCCCCA  613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATCGAGATCTGCGGCCATAAGGCCATCGGCACCGTGCTGGTCGGCCCCCA 3340

Qy    614 CCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCACCCTGAACTTCC  673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCGTGAACATCATCGGCCGCAACATCGGCTGACTGCTGACCCAGATCGGCTGCACCCTGAACTTCC 3400

Qy    674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCTGGAAGCCCGCATGACGGCCCCAAGG  733
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCTGGAAGCCCGCATGACGGCCCCAAGG 3460

Qy    734 TGAAGCAGTGCCCCTGACCAGCGACGAGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA  793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGCCCCTGACCAGCGACGAGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 3520

Qy    794 TGGAAGAAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACACACACCCCGTGT  853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGAAGAAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACACACACCCCGTGT 3580

Qy    854 TCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGGCAGCTGA  913
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGGCAGCTGA 3640

Qy    914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCGGCCCTGA  973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCGGCCCTGA 3700

Qy    974 AGAGAAGAGAGCGTGACCGTGTGGACGTGTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 1033
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAGAAGAGAGCGTGACCGTGTGGACGTGTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGG 3760

Qy   1034 ACGAGGACTTCCGCAAGTACACCGCGCTTCACCATCCCCAGGGCTGGCATCATCAACAACGAGACCCCCG 1093
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACGAGGACTTCCGCAAGTACACCGCGCTTCACCATCCCCAGGGCTGGCATCATCAACAACGAGACCCCCG 3820

Qy   1094 GCATCCGTACCAGTACAACGTGCTGCCCCCAGGGCTGAAGGGCAGCCCCAGCATCTTCC 1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGTACCAGTACAACGTGCTGCCCCCAGGGCTGAAGGGCAGCCCCAGCATCTTCC 3880

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACAGCCCGAGATCGTGATCT 1213
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3881 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACAGCCCGAGATCGTGATCT 3940

Qy   1214 ACCAGGCCCCCCTGTACGTGGGCCAGCCAGCCAGATCGGCGCCAAGATCG 1273
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3941 ACCAGGCCCCCCTGTACGTGGGCCAGCCAGCCAGATCGGCGCCAAGATCG 4000

Qy   1274 AGGAGTGCGGCAAGCACCTGCTGCCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 1333
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4001 AGGAGTGCGGCAAGCACCACGTGCTGCCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 4060

Qy   1334 AGGAGGCCCCCTTCCTGTGGATGGGCTACAGAGTGGCTACGAGTGGACCGTGCAGC 1393
          |||||||||||||||||||          |||||||||||||||||||||||||||
Db   4061 AGGAGGCCCCCTTCCTGTGCCCAT------CGAGCTGACCCCGACAAGTGGACCGTGCAGC 4114

Qy   1394 CCATCGAGCTGCCCGAGAAGGAGAGCTGAACCGTGAACACATCCAGAAGCTGGTGGGCA 1453
          |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   4115 CCATCGAGCTGCCCGAGAAGGAGAGCTGAACGACATCCAGAAGCTGGTGGGCA 4174

Qy   1454 AGCTGAACTGGGCCAGATCTACCCCGGCATCAAGGTGCGCCAGTCGTGCAAGCTGC 1513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4175 AGCTGAACTGGGCCAGATCTACCCCGGCATCAAGGTGCGCCAGTCGTGCAAGCTGC 4234

Qy   1514 TGCGCGGCGCCAAGGCCCTGACCGACGTGCCCCTGACCGAGGAGATGAGCCTGGAGC 1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4235 TGCGCGGCGCCAAGGGCCCTGACCGACGTGCCCCTGACCGAGGAGATGAGCCTGGAGC 4294

Qy   1574 TGGCCCGAGAACCGCGAGATCCTGCGGCGGCGTGTACGACCCCAGCA 1633
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   4295 TGCCCGAGAACCGCGAGATCCTGCGGCGGCGTGTACGACCCCAGCA 4354

Qy   1634 AGGACCTGGTGGCCGAGATCCAGGAGGCCAGTCAGTGGACCTACCAGATCTACC 1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4355 AGGACCTGGTGGCCGAGATCCAGGAGGCCAGTCAGTGGACCTACCAGATCTACC 4414

Qy   1694 AGGAGCCCTTCAAGAACCTGAAGAACCTGAGAACGGCCAAGATGCCCAAGATGCCCACCACCA 1753
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4415 AGGAGCCCTTCAAGAACCTGAAGAACCTGAGAACGGCCAAGATGCCCAAGATGCCCACCACCA 4474

Qy   1754 ACGACGTGAAGCAGCAGTCAGAGAGCTGACCCGTGGGCTGGGCTGCAGAAGATCGTGATCT 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4475 ACGACGTGAAGCAGCAGTCAGAGAGCTGACCCGTGGGCTGGGCTGCAGAAGATCGTGATCT 4534
```

Matches 2414; Conservative 0; Mismatches 24; Indels 6; Gaps 1;

```
Qy    14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCGCCAACTCCTGATGCGAGCAGCAACTTCA 73
         |||||||||||||||||||||  ||| |||
Db  1088 TGGCCGAGGCGATGAGCCAGGCCAACACCAGCGGTGATGATGCGAGCAGCAACTTTAAAA 1147

Qy    74 AGGGCCCCAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGCGGCCACATCGCCGGCA 133
         |||||||||||||||||||| ||||||||||||||||||||||||||
Db  1148 AGGGCCCCAAGCGCATCAAGTGCTTCAACTGCGGCAAGGAGCGGCCACATCGCCGGCA 1207

Qy   134 ACTGCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGTGGCCACCCAGATGA 193
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1208 ACTGCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGTGGCCACCCAGATGA 1267

Qy   194 AGGACTGCGACCGAGGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCCA 253
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1268 AGGACTGCACCGAGGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCCA 1327

Qy   254 AGGGCCCCGGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCAGCCGCGGCGAGCTGC 313
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1328 AGGGCCCCGGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCAGCCGCGGCGAGCTGC 1387

Qy   314 AGGTGCGGCGGCACACCCCGAGCGGCGAGCGGCCCAGCGGCGCACCCTGAACT 373
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1388 AGGTGCGGCGGCACAACCCCGAGCGGCGAGCGGCCCAGCGGCGCACCCTGAACT 1447

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGGTGAGCATCAGCCAGCCAGATCATCG 433
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1448 TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGGTGAGCATCAGCCAGCCAGATCATCG 1507

Qy   434 AGGAGGCCCTGCTGGACACCGGCCGCCCACAACCGTGGTGGGAGATGAGCCTGCCCG 493
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1508 AGGAGGCCCTGCTGGACACCGGCCGCCCACAACCGTGGTGGGAGATGAGCCTGCCCG 1567

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGCATCGGCGCCAAGAAGGGCCATCGGAGCCAGTACG 553
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1568 GCAAGTGGAAGCCCAAGATGATCGGCGCATCGGCGCCAAGAAGGGCCATCGGAGCCAGTACG 1627

Qy   554 ACCAGATCCTGATCGAGATCTGCGGCCAACATGCTGACCCAGTGGGCGTGCACCCTGAACTTCC 613
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1628 ACCAGATCCTGATCGAGATCTGCGGCCAACATGCTGACCCAGTGGGCGTGCACCCTGAACTTCC 1687

Qy   614 CCCCCGTGACACATCATCGGCGCCGCAACATGCTGCCCGTGAAGCTGAAGCCCGGCATGGACGG 673
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1688 CCCCCGTGAACATCATCGGCGCCGCAACATGCTGCCCGTGAAGCTGAAGCCCGGCATGGACGG 1747

Qy   674 CCATCAGCCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCGACGGCCCCAAGG 733
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1748 CCATCAGCCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCTGGAGCTGAAGCCCGGCGACGGCCCCAAGG 1807

Qy   734 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGGCCATCTGCGAGGAGA 793
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1808 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGGCCATCTGCGAGGAGA 1867

Qy   794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACACCCTACAACACCCCGTGT 853
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1868 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACACCCTACAACACCCCGTGT 1927

Qy   854 TCGCCATCAAGAAGGAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGGGAGCTGA 913
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1928 TCGCCATCAAGAAGGAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGGGAGCTGA 1987

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCCGCCGGCCTGA 973
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1988 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCCGCCGGCCTGA 2047

Qy   974 AGAAGAAGAAGAGCGTGACGCTGCTGGACGTGGGCGACGCCCTACTTCAGCCGGCCCCTGG 1033
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2048 AGAAGAAGAAGAGCGTGACGCTGCTGGACGTGGGCGACGCCCTACTTCAGCCGGTGCCCCTGG 2107

Qy  1034 ACGAGGACTTCCGCAAGTACACCCCG 1093
         |||||||||||||||||||||||||
Db  2108 ACGAGGACTTCCGCAAGTACACCCCG 2167

Qy  1814 GGGGCAAGACCCCCAAGTTCCGCCTGCCCTGCCCATCCAGAAGGAGGAGACCTGGGAGACCTGGTGGA 1873
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4535 GGGGCAAGACCCCCAAGTTCCGCCTGCCCTGCCCATCCAGAAGGAGGAGACCTGGGAGACCTGGTGGA 4594

Qy  1874 CCGACTACTGGGAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCCCTGG 1933
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4595 CCGACTACTGGGAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCCCTGG 4654

Qy  1934 TGAAGTGTGGTACCAGCTGGAGAAGGAGACCCCATCATCGGCGGCAAGGAGACCTTCTACGTGG 1993
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4655 TGAAGTGTGGTACCAGCTGGAGAAGGAGACCGCCATCATCGGCGGCAAGGAGACCTTCTACGTGG 4714

Qy  1994 ACGGGCGCGCCAACCGGCAGAACCCGCAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGCC 2053
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4715 ACGGCGCGCCAACCGGCAGAACCCGCAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGCC 4774

Qy  2054 GGCAGAAGATCGTGAGCGCTGCACCGAGACCACCAACCAGAAGACCGAGCTGCAGGGCCATCC 2113
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4775 GGCAGAAGATCGTGAGCGCTGCACCGAGACCACCAACCAGAAGACCGAGCTGCAGGGCCATCC 4834

Qy  2114 AGCTGGCCCTGCAGGACAGCGGCCAGCGGCGAGGTGAACATCGTGACCGACAGCCAGTACGCGCCC 2173
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4835 AGCTGGCCCTGCAGGACAGCGGCCAGCGGCGAGGTGAACATCGTGACCGACAGCCAGTACGCGCCC 4894

Qy  2174 TGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCG 2233
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4895 TGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCG 4954

Qy  2234 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG 2293
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4955 AGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCG 5014

Qy  2294 GCGCGCAACGAGCAGAGATGCGACAAGCTGGTGAGACAAGGGCATCCGCCAAGGTGCGTGTTCCTGG 2353
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  5015 GCGCGCAACGAGCAGCAGATGCGACAAGCTGGTGAGACAAGGGCATCCGCCAAGGTGCGTGTTCCTGG 5074

Qy  2354 ACGGCATCGATGGCGGCATCGTGATCTACCAGTACACATGGACGACCTGTACGTGGGCAGCG 2413
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  5075 GCGGCATCGATGGCGGCATCGTGATCTACCAGTACACATGGACGACCTGTACGTGGGCAGCG 5134

Qy  2414 GCGGCCCCTAGGATCGATTAAAAAGCTTCCGGGGGCTAGCACCGGT 2457
         ||||||||||||||||||||||||||||||||||||||||||
Db  5135 GCGGCCCTAGGATCGATTAAAAAGCTTCCGGGGGCTAGCACCGGT 5178
```

RESULT 12
US-10-190-435-13
; Sequence 13, Application US/10190435
; Publication No. US20030143248A1
GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 3531
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagPolmut_C
US-10-190-435-13

Query Match   96.8%; Score 2383.6; DB 14; Length 3531;
Best Local Similarity   98.8%; Pred. No. 0;

```
Qy   2174  TGGGCATCATCCAGGCCCAGCCCGACACAGAGCGAGAGACGGAGCTGGTGAACCAGATCATCG  2233
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3242  TGGGCATCATCCAGGCCCAGCCCGACACAGAGCGAGAGACGGAGCTGGTGAACCAGATCATCG  3301

Qy   2234  AGCAGCTGACTCAAGAAGGAGAAGGTGTACCTGAGTCGACTCGAGTCGTGCCCCACAAGGGCATCG  2293
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3302  AGCAGCTGACTCAAGAAGGAGAAGGTGTACCTGAGTCGACTCGAGTCGTGCCCCACAAGGGCATCG  3361

Qy   2294  GCGGCAACGAGCAGATCACCAAGCTGGTGAGCAAGGCATCCGCAAGTGTCGTGTTCCTGG  2353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3362  GCGGCAACGAGCAGATCACCAAGCTGGTGAGCAAGGCATCCGCAAGTGTCGTGTTCCTGG  3421

Qy   2354  ACGGGCCCTAGACGTGGCGGCCATCGTGATCTACCAGTACATCGAGCACCTGTACGTGGGCAGCG  2413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3422  ACGGGCCCTAGACGTGGCGGCCATCGTGATCTACCAGTACATCGAGCACCTGTACGTGGGCAGCG  3481

Qy   2414  GCGGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGT  2457
           |||||||||||||||||||||||||||||||||||||||
Db   3482  GCGGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGT  3525
```

RESULT 13

US-10-190-435-14

; Sequence 14, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Estrelita J.

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: PP18133.003 / 2302-18133

; CURRENT APPLICATION NUMBER: US/10/190,435

; CURRENT FILING DATE: 2002-12-30

; NUMBER OF SEQ ID NOS: 319

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14

; LENGTH: 3537

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: GagPolmutatt_C

US-10-190-435-14

```
Query Match     96.7%;  Score 2382;   DB 14;   Length 3537;
Best Local Similarity  98.7%;   Pred. No. 0;
Matches 2413;   Conservative   0;  Mismatches   25;   Indels   6;   Gaps    1;
```

```
Qy     14  TGGCCGAGGCCATGAGCCAGGCCACCAGGCCCAACATCCTGATGCAGCGCAGCAACTTCA    73
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1094  TGGCCGAGGCGATGAGCCAGGCCAGGCCAACACCAGCGTGATGATGCAGAGAGAGCAACTTTAAAA  1153

Qy     74  AGGGCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA   133
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1154  AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCCCGCA  1213

Qy    134  ACTGCCGGGCGCCCCCGCAAGGAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA   193
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1214  ACTGCCGGGCGCCCCCGCAAGGAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA  1273

Qy    194  AGGACTGCACCGAGCGCCAACTTCTTCCGGCAGCTGGCCTTCCCCCAGGCA   253
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274  AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGCAGCTGGCCTTCCCCCAGGCA  1333

Qy    254  AGGGCCCGCGAGTTCCCAGCGAGACCGAGAACCGGGCCAACAGCCGCGAGTCGC   313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334  AGGGCCCGCGAGTTCCCAGCGAGACCGAGAACCGGGCCAACAGCCGCGAGTCGC  1393

Qy    314  AGGTGCGCGGCGACAACCCCGCGCGAGCCCCGCGAGGGGCCACCCTGACT   373
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1394  AGGTCGCGCGGCGACAACCCCGCGCGAGCCCGCGCGAGGGGCCACCCTGACT  1453
```

```
Qy   1094  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGGCAGCGCCCCAGCATCTTCC  1153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2168  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGGCAGCGCCCCAGCATCTTCC  2227

Qy   1154  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGGCCCAGAACCCCGAGATCGTGATCT  1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2228  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGGCCCAGAACCCCGAGATCGTGATCT  2287

Qy   1214  ACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGGGCCAAGATCG  1273
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2288  ACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGGCCAGCACCGGGCCAAGATCG  2347

Qy   1274  AGGAGTGGCAAGCACCTGCGCGTGGGGGCTTCACCACCCCGACAAGAGCACCAGA  1333
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2348  AGGAGTGGCGCAAGCACCTGCGCGTGGGGGCTTCACCACCCCGACAAGAGCACCAGA  2407

Qy   1334  AGGAGCCCCCTTCCTGTGGATGGGCTACGACGTGCACCCGCTGCAGC  1393
           |||||||||||||||||||||||||||||||||||||||
Db   2408  AGGAGCCCCCTTCCTGTGCGCTGGGCTGCGGGCTGGGCTGGGGCA  CGABTCGTGCACCCGCTGCAGC  2461

Qy   1394  CCATCGAGCTGCCCGAGAAGGAGAGTGACCGTGAACGACATCCAGAAGCTGGTGGGCA  1453
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2462  CCATCGAGCTGCCCGAGAAGGAGAGTGACCGTGAACGACATCCAGAAGCTGGTGGGCA  2521

Qy   1454  AGCTGAACTGGGCCAGCGAGATCCTGCGCCAGCTGTGCGCCAGCTGTGCCAGCTGC  1513
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2522  AGCTGAACTGGGCCAGCGAGATCTACCCCGGCCATCAACAAGGTGCGCCAGCTGTGCAAGCTGC  2581

Qy   1514  TGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGGAGGCCGAGCTGAGC  1573
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2582  TGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGGAGGCCGAGCTGAGC  2641

Qy   1574  TGGCCGAGAACCGCGAGATCCTCGCGCAGCCCGTGTACTGCGGCGGTGTACTGCGACCCCAGCA  1633
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2642  TGGCCGAGAACCGCGAGATCCTGCGCGCAGCCCGTGTACTGCGGGCGGTGTACTGCGACCCCAGCA  2701

Qy   1634  AGGACCTGGTTGCCGAGATCCAGAAGCAGAGGGGCCACGACCAGTTGGACCTTACCAGATCTACC  1693
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2702  AGGACCTGGTTGCCGAGATCCAGAAGCAGAGGGGCCACGACCAGTTGGACCTTGACCTACCAGATCTACC  2761

Qy   1694  AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCCACCGCCCCACACCA  1753
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2762  AGGAGCCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCCACCGCCCCACCA  2821

Qy   1754  ACGACGTGAAGCAGCTGACCGAAGCCGTGCAGAAGATCGCATGGGAGCATCGTGATCT  1813
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2822  ACGACGTGAAGCAGCTGACCGAAGCCGTGCAGAAGATCGCCATGGGAGAGCATCGTGATCT  2881

Qy   1814  GGGGCAAGACCCCAGTTCCGCCTGCCCATCGAGAAGGACCTGGGAGACCTGGTTGGA  1873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2882  GGGGCAAGACCCCCAGTTCCGCCTGCCCATCGAGAAGGACCTGGGAGACCTGGTGGA  2941

Qy   1874  CCGACTACTGGCAGGCCACCTGGATCCCCGAGTTGGAGTTCGTGAACACCCCCCCCCTGG  1933
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2942  CCGACTACTGGCAGGCCACCTGGATCCCCGAGTTGGGAGTTCGTGAACACCCCCCCCCTGG  3001

Qy   1934  TGAAGCTGTGTGGTACCAGTGGTGGAGAAGGGAGAAGGAGCCCCATCATCGGGCGCCAGTCGTGG  1993
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3002  TGAAGCTGTGTGTACCAGTGGTGGAGAAGGAGAAGGAGCCCCATCATCGGCGCCAGTCGTGG  3061

Qy   1994  ACGGGCCGCCAACCGCGAGATCGGCAAGGCCGCTACGTGACCGACCGGGGCC  2053
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3062  ACGGGCCGCCAACCGCGAGATCGGCAAGGCCGCTACGTGACCGACCGGGGGCC  3121

Qy   2054  GGCAGAAGATCGTGAGCCTGGAGCCTGACCGAGAACCAACCAGAAGACCGAGTCTGCAGGCCATCC  2113
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3122  GGCAGAAGATCGTGAGCCTGGAGCCTGACCGAGAACCAACCAGAAGACCGAGTCTGCAGGCCATCC  3181

Qy   2114  AGCTGGCCCTGCAGGACAGCGGACTGGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCC  2173
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3182  AGCTGGCCCTGCAGGACAGCGGACTGGCGAGGTGAACATCGTGACCGACAGCCAGTACGCGCCC  3241
```

```
Qy    374  TCCCCCAGATCACCCTGTGTGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA   433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1454  TCCCCCAGATCACCCTGTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA  1513

Qy    434  AGGAGGCCCCTGCTGACACCGGCGCCGACGACACCGGCGCCGACGACACTGCTGCCGG     493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1514  AGGAGGCCCCTGCTGACTCCGGCGCCGACGACTCCGGCGCCGACGATGAGCCTGCCCG    1573

Qy    494  GCAAGTGGAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG    553
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1574  GCAAGTGGAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG   1633

Qy   1574  ACCAGATCTGATCGAGATCTGCGGCGCAAGAAGGCCATCGGCGACAAGGTGCTGATCGGCCCA   613
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1634  ACCAGATCCTGATCGAGATCGCGGCGCAAGAAGGCCATCGGCGACAAGGTGCTGATCGGCCCCA   1693

Qy    614  CCCCCGTGAACATCATCGGCGCCGACATGCTGCCCAGCTGGCGGCTGCACCCTGAACTTCC    673
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1694  CCCCCGTGAACATCATCGGCGCCGACATGCTGCCCAGCTGGCGGCTGCACCCTGAACTTCC   1753

Qy    674  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCATCTGCGAGGAGA    733
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1754  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCATCTGCGAGGAGA   1813

Qy    734  TGAAGCAGTGGCCCCTGACCGAGGAGAATCAAGGCCCTGACTTCCGCGAGGAGA    793
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1814  TGAAGCAGTGGCCCCTGACCGAGGAGAATCAAGGCCCTGACTTCCGCGAGGAGA   1873

Qy    794  TGGAGAGGGAGGGCAAGATCACCAGACTTCTGGGAGGTGCGCAGCTGGCATCCCCGTGT    853
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1874  TGGAGAGGGAGGGCAAGATCACCAGACTTCTGGGAGGTGCGCAGCTGGCATCCCCGTGT   1933

Qy    854  TCGCCATCAAGAAGTGGCCCGAGGAGAGCACCAGACTGGCGCAAGGTCGTGTGACTTCCGCGAGCTGA    913
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1934  TCGCCATCAAGAAGTGGCCCGAGGAGAGCACCAGACTGGCGCAAGGTCGTGTGACTTCCGCGAGCTGA   1993

Qy    914  ACAAGGCCACCCAGCACTTCTGGGAGTGCGAGCTGCGCATCCCCGGCCGCCCGGCCTGA    973
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1994  ACAAGGCCACCCAGCACTTCTGGGAGTGCGAGCTGCGCATCCCCGGCCGCCCGGCCTGA   2053

Qy    974  AGAAGAAGAGCGTGCAGCGTGACCGGTGGGCACGCGCCGCCCGCCAGCGCCATCGTGTCC   1033
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2054  AGAAGAAGAGCGTGCAGCGTGACCGGTGGGCAGCGCCCGCCCGCCAGCGCCATCGTGTCC   2113

Qy   2054  AGCAGCAGATGACCAAGATCCTGAGCGTGGGGCAGCCCGGACCCCTTCCGCGCCCCCCG   2113
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1034  AGCAGCAGATGACCAAGATCCTGAGCGTGGGGCAGCCCGGACCCCTTCCGCGCCCCCCG   1093

Qy   2114  AGCAGGACTTCGCGCAAGTACACCGCCTTCACCCATCCCAGCATCAACGAACGAGACCCCG   2173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1094  AGCAGGACTTCGCGCAAGTACACCGCCTTCACCATCCCAGCATCAACGAACGAGACCCCG   1153

Qy   1154  GCATCCGCTACCAGTACTACAACGTGCTGCTGGGCGACGCGCTACTTCAGGGCTGCCCCTGG   1213
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2174  GCATCCGCTACCAGTACTACAACGTGCTGCTGGGCGACGCGCTACTTCAGGGCTGCCCCTGG   2233

Qy   1214  AGAGCAGATGACCAAGATCCTGAGCGTGGGGCAGCCCTTCCGCGGCCTTCCGCGAGATCT   1273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2234  AGAGCAGATGACCAAGATCCTGAGCGTGGGGCAGCCCTTCCGCGGCCTTCCGCGAGATCT   2293

Qy   2294  GCGGGCAACGCCCCCCCCGGACGACCCCGCCCGCCAGCGCCCAGCGCCCCAGCATCAACGAGA   2353
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1274  GCGGGCAACGCCCCCCCCGGACGACCCCGCCCGCCAGCATCAACGAGA   1333

Qy   1334  ACCAGGACCCCCTTCCTGTACGTGGATGGGCTACGAGCTGACAGTTGCACCGTGCAGC   1393
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2354  ACCAGGACCCCCTTCCTGTACGTGGGGCAGCCTACGAGCTGACAGTTGCACCGTGCAGC   2413

Qy   2414  GCATTCGCTACCAGTACTACAGTCCGTGACCGTGAACATCAGCATCCAGAAGACGCACCAGA   2467
           |||||||||||||||||||||||||||||||||||||||||||
Db   1334  GCATCCGCTACCAGTACTACAGTCCGTGACCGTGAACATCAGCATCCAGAAGACGCACCAGA   1393

Qy   1394  CCATCGAGCTGCCCGAGAAGGAGAGGCTGCCCGTGAACACATCCAGAAGATCCAAGAAGTCGTGGGCA   1453
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2468  CCATCGAGCTGCCCGAGAAGGAGAGCTGCCCGTGAACGACATCAGCATCCAAGAAGTCGTGGGCA   2527
```

```
Qy   1454  AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGTCGCAAGCTGC   1513
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2528  AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGTCGCAAGCTGC   2587

Qy   1514  TGCGCGGCGCCAAGGCCCTGACCGACCATCGTGCCCCTGACCGAGGAGGCCGAGCTGGGAGC   1573
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2588  TGCGCGGCGCCAAGGGCCCTGACCGACCATCGTGCCCCTGACCGAGGAGGCCGAGCTGGGAGC   2647

Qy   1574  TGGCCGAGAACCGGCGAGATCCTGCGCGACAGCCCGTGCACGGCGGCGTGTACTACGACCCCAGCA   1633
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2648  TGGCCGAGAACCGGCGAGATCCTGCGCGACAGCCCGTGCACGGCGGCGTGTACTACGACCCCCAGCA   2707

Qy   1634  AGGACCTGGTGGCCGAGATCCAGAACGCAGGGCCACGACCAGTGACCTACCAGATCTACC   1693
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2708  AGGACCTGGTGGCCGAGATCCAGAACGCAGGGCCACGACCAGTGACCTACCAGATCTACC   2767

Qy   1694  AGGAGCCCTTCAAGAACCTGAAGAACCGGCAAGTACGGCCAAGTACGGCGCCACACCA   1753
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2768  AGGAGCCCTTCAAGAACCTGAAGAACCGGCAAGTACGGCCAAGTACGGCGCCACACCA   2827

Qy   1754  ACGACGTGAAGCAGCTGACGAGCTGCTGACCGGCCGTGCGAGAAGATCGTGTGATCT   1813
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2828  ACGACGTGAAGCAGCTGACGAGCTGCTGACCGGCCGTGCGAGAAGATCGTGTGATCT   2887

Qy   1814  GGGGCAAGACCCCCAAGTTCCGGCTGCTGCCCATCCCAGAAGGAGACCTGGAGACCTGTGTGGA   1873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2888  GGGGCAAGACCCCCAAGTTCCGGCTGCTGCCCATCCCAGAAGGAGACCTGGAGACCTGTGTGGA   2947

Qy   1874  CCGACTACTGGCAGGCCACCTGGATCCCCGGATCCCCGAGTGGGAGGTTCGTGAACACCCCCTGG   1933
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2948  CCGACTACTGGTGTACCAGCTGGCAGGCCACCTGGATCCCCGAGTGGGAGGTTCGTGAACACCCCCTGG   3007

Qy   1934  TGAAGCTGTGTACCAGTGGTACCAGCTGGAGAAGAGCCCATCATCGGCGCCCGAGATCCTGG   1993
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3008  TGAAGCTGTGTACCAGTGGTACCAGCTGGAGAAGAGCCCATCATCGGCGCCCGAGATCCTGG   3067

Qy   1994  ACGCGGGCGCCAACCGCGAGACCTGAGCGTTCACCGACGAGCTGCAGGCCATCC   2053
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3068  ACGCGGGCGCCAACCGCGAGACCTGAGCGTTCACCGACGAGCTGCAGGCGGGGCC   3127

Qy   2054  GGCAGAAGATCGTGAGCGTGCTGCTGAGCGTGACCGCCGAGGCGCAGACCAGATCATCG   2113
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3128  GGCAGAAGATCGTGAGCGTGCTGCTGAGCGTGACCGCCGAGGCGCAGACCAGATCATCG   3187

Qy   2114  AGCTGCCCTGCAGACAGCGGCAGCCCGACAGCGGAAATCGTGACCGAGGTGACATCGCCC   2173
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3188  AGCTGCCCTGCAGACAGCGGCAGCCCGACAGCGGAAATCGTGACCGAGGTGACATCGCCC   3247

Qy   2174  TGGGGCATCATCCAGGCCGTGCCCCCAGGGCTGAAGGCTGAAGGCTGGAACCAGAGATCATCG   2233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3248  TGGGGCATCATCCAGGCCGTGCCCCCAGGGCTGAAGGCTGAAGGCTGGAACCAGAGATCATCG   3307

Qy   2234  AGCAGCGATGGCGGCATCGGCGGCCTTCGGCGCTCGGGCTGAGTGTACCTGAGCGTGTACCTGAGT   2293
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3308  AGCAGCGATGGCGGCATCGGCGGCGTACTGAGTGTACCTGAGCGTGTACCTGAGT   3367

Qy   2294  GCGGGCAACGAGCAGCTGATCGAACAGCTGGTGAGCGAACAAGGTGGTGACGGCATCCGCGAAGGGCATCG   2353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3368  GCGGGCAACGAGCAGCTGATCGAACAGCTGGTGAGCGAACAAGGTGGTGACGGCATCCGCGAAGGGCATCG   3427

Qy   2354  ACGGCATCGATGGCGGCATCGGCGGCGTTCACCAGTCAATCGTGATCTACCAGTACCATGACACGACGGGCATCG   2413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3428  ACGGCATCGATGGCGGCATCGGCGGCGTGATCTACCAGTACCATGACACGACGGGCAGCG   3487

Qy   2414  GCGGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT   2457
           |||||||||||||||||||||||||||||||
Db   3488  GCGGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGT   3531
```

RESULT 14
US-10-190-435-15
; Sequence 15, Application US/10190435
; Publication No. US20030143248A1

```
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 15
;   LENGTH: 3537
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: GagPolmutIna_C
US-10-190-435-15
```

```
Query Match         96.7%;   Score 2382;  DB 14;  Length 3537;
Best Local Similarity  98.7%;   Pred. No. 0;
Matches 2413;  Conservative  0;  Mismatches  25;  Indels  6;  Gaps  1;

Qy        14 TGGCCAGGAGGCCATGAGCCAGGCCACCAGCGCCAGCATCCTGATGCAGCGGCAGCAACTTCA 73
             ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1094 TGGCCGGAGGCGATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGCAACTTTAAAA 1153

Qy        74 AGGGCCCCAAGGCATCATCAAGTGTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 133
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1154 AGGGCCCCAAGGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 1213

Qy       134 ACTGCCGCGCCCCCCGCAAGAAGGGCGTGTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1214 ACTGCCGCGCCCCCCGCAAGAAGGGCGTGCGGAAGTGCGGCAGCGAGGGCCACCAGATGA 1273

Qy       194 AGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGCGAAGACCTGGCCTTCCCCAGGGCA 253
             |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      1274 AGGACTGCACCGAGCGGCAGGCCAACCTGCCTTCACCGAGACCTGGCCTTCCCCAGGGCA 1333

Qy       254 AGGCCCGCGAGTTCCCCAGCCGGAGAGAACCGCGCCAACAGCGCCCACCAGCGCGGAGCTGC 313
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1334 AGGCCCGCGAGTTCCCCAGCCGGAGAGAACAGCGCCCACCAGCGCGGGAGCTGC 1393

Qy       314 AGGTGGCGGGGACAACCCCGCAGCGAGGCCGGAGCGCCCAGCGGCCAGGGCCACCCTGAACT 373
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1394 AGGTGCGCGGGCAGACAACCCCCGCAGCGAGGCCGGAGCGCCCAGGGCCACCCTGAACT 1453

Qy       374 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGTGTGAGCATCAAGGTGGCGGCCAGATCA 433
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1454 TCCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCCAGATCA 1513

Qy       434 AGGAGGCCCTGCTGGACAACCGGCCGCGGCCGGCCGACAACCGTGCTGGAGGATGAGCCTGCCCG 493
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1514 AGGAGGCCCTGGTGGCCGAGATCCGCGGCCAAGAGAGGCCATCGCCACCGTGCTGCTGCCCG 1573

Qy       494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1574 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 1633

Qy       554 ACCAGATCCTGATCGAGATCTGCGGCGAAGAGGCCATCGCACCGTCTGCTGATCGGCCCG 613
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1634 ACCAGATCCTGATCGAGATCTGCGGCGAAGAGGCCATCGCACCGTCTGCTGATCGGCGCCCCA 1693

Qy       614 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1694 CCCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 1753

Qy       674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAGG 733
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1754 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCCAGG 1813
```

```
Qy       734 TGAAGCAGTGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGAGA 793
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1814 TGAAGCAGTGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGAGA 1873

Qy       794 TGGGAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACACACCCCCGTGT 853
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1874 TGGGAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACACACACCCCGTGT 1933

Qy       854 TCGCCATCAAGAAGAAGGACACCAGCAAGTGGCGCAAGCTGGTGGACTTCCGGGACTGA 913
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1934 TCGCCATCAAGAAGAAGGACAGCAAGTGGCGCAAGCTGGTGGACTTCCGGGAGCTGA 1993

Qy       914 ACAAGGCGCACCCAGGACTTCTGGGAGGTGGCAGCTGGCATCCCCCACCCGCCGGCCTGA 973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1994 ACAAGGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCATCCCCCACCCGCCGGCCTGA 2053

Qy       974 AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGAGCTGGCCTACTTCAGCGTGCCCTGG 1033
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2054 AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGACGTGCTACTTCAGCGTGCCCTGG 2113

Qy      1034 ACGAGGACTTCCGCAAGTACACCACCGCCTTCACCATCCCAGCATCAACAACGAGACCCCG 1093
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2114 ACGAGGACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCG 2173

Qy      1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 1153
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2174 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 2233

Qy      1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 1213
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2234 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 2293

Qy      1214 ACCAGGGCCCCCTGTACGTGGGCCAGCGACCTGGAGCTGGGCCAGCGGCCAGGATCG 1273
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2294 ACCAGGGCCCCCTGTACGTGGGCCAGCGACCTGGAGCTGGGCCAGCGGCCAGGATCG 2353

Qy      1274 AGGACTGCGCAAGCACCTGCTGCGCTGGGCTGGCCTTCACCACCCCGACAAGAAGACCACCAGA 1333
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2354 AGGAGTGCGCAAGCACCTGCTGCGCTGGGCCTTCACCACCCCGACAAGAAGACCACCAGA 2413

Qy      1334 AGGACGCCCCTTCCTGTGAAGCGCTACGAGCTGCACCCGACAAGGTGGACCGTGCGAGC 1393
             ||||||||||||||||||||||              |||||||||||||||||||
Db      2414 AGGAGCGCCCCTTCCTTCCTGCCCAT------CGAGCTGCACCCGACCAAGTGGACCGTGCGAGC 2467

Qy      1394 CCATGCGAGCTGCCCGAGAAGGAGAGCTGACCCGGATCCAGAAGCATCCAGAAGCTCGTGGTGGCA 1453
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2468 CCATCGAGCTGCCCGAGAAGGAGAGCTGACCCGGATCCAGAACCACATCCAGAAGCTGTGTGGGGA 2527

Qy      1454 AGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGTGCGGCCAGATCA 1513
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2528 AGCTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGGTGCCCAGCGTGTGCAAGTTGC 2587

Qy      1514 TGCGCGGCGCCAAGGCCCTGACCGACACCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGC 1573
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2588 TGCGCGGCGCCAAGGCCCTGACCGACACCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGC 2647

Qy      1574 TGGCCGAGAACCGGCGGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCA 1633
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2648 TGGCCGAGAACCGGCGGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCA 2707

Qy      1634 AGGACCTGGTTGGCCAGATCCAGAAGCAGGGCCACGACCAGCAGGCCACCAGATCTACC 1693
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2708 AGGACCTGGTTGGCCAGATCCAGAAGCAGGGCCACGACCAGCAGGCCACCAGATCTACC 2767

Qy      1694 AGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGCACACCGGCCACACCA 1753
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2768 AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGCACACCGGCCACACCA 2827

Qy      1754 ACAGCGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATCGAGAGCATCGTGATCT 1813
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2828 ACAGCGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATCGAGAGCATCGTGATCT 2887

Ov      1814 GGGGCAAGCACCCCCAAGTTCGCGCCTGCCCATCCAGAAGAGAGACCTGGGAGACCTGGTGA 1873
```

```
Db    2888  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
            GGGGCAAGAGACCCCCAAGTTCCGCCTGCCATCCAGAAGGAGGACACCTGGTGGGA  2947
Qy    1874  CCGACTACTGGCAGGCCACCTGGATCCCCGAGTTGGGAGTTCGTGAACACCCCCCTGG  1933
Db    2948  CCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGG  3007
Qy    1934  TGAAGCTGTGTGTACCAGCTGGAGAAGGAGATC  1993
Db    3008  TGAAGCTGTGTGTACCAGCTGGAGAAGGAGATC  3067
```

```
Qy    14    TGGCCGAGGCCATGAGCAGCCAGCGCCACCAGCGCCAACATCCTGATGAGCGAGCAACTTCA  73
Db    1487  TCGCCGAGGCCATGAGCAGCCAGCGCCACCAGCGCCAACATCCTGATGAGCGAGCAACTTCA  1546
Qy    74    AGGGCCCAAGGCGCATCATCAAGTGCTTCAACTGGCGCAAGAGAGGGCCACATCGCCCGCA  133
Db    1547  AGGGCCCAAGGCGCATCATCAAGTGCTTCAACTGGCGCAAGAGAGGGCCACATCGCCCGCA  1606
Qy    134   ACTCCGCGCCCCCGCGCCCCGCGCCAAGAAGGGTGCTGGAAGTGCGGCAAGAGGGCCACCAGATGA  193
Db    1607  ACTCGCGCCGCCCCGCGCCCCGCGCCAAGAAGGGCTCTGGAAGTGCGGCAAGAGGGCCACCAGATGA  1666
Qy    194   AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGGAGGACCTGGCCTTCCCCCAGGCA  253
Db    1667  AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGGAGGACCTGGCCTTCCCCCAGGGCA  1726
Qy    254   AGGGCCGGCGAGTTCCCAGCGAGCAGAACCGGCCCAACAGCCCCACGGCGGCGAGCTGC  313
Db    1727  AGGGCCGGCGAGTTCCCAGCGACGAGAACCGGCCCAACAGCCCCACGGCGGCGAGCTGC  1786
Qy    314   AGGTGCGGCGGCGACAACCCCCCAGCGAGGCGCCCCCGCGGGG  373
Db    1787  AGGTGCGGCGGCGACAACCCCCCAGCGAGGCGCCCCCGCGGGG  1846
Qy    374   TCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCGAAGGTGGCGGCGCCAGATCG  433
Db    1847  TCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCGCCAGATCA  1906
Qy    434   AGGAGGCCCTGCTGGACACCGGCCGCCGCCGCGACGACACCTGGCGCTTCATCGGCACCGTG  493
Db    1907  AGGAGGCCCTGCTGGCACCGGCCGCCGCCGACGACACCTGGCGCTTCATCGGCACCGTG  1966
Qy    494   GCAAGTGGAAGCCCAAGATGATCGGCCGCAAGAAGCCATCGGCACCCGTGCTGATCGGCCCA  553
Db    1967  GCAAGTGGAAGCCCAAGATGATCGGCCGCAAGAAGCCATCGGCACCCGTCGTCGTGATCGGCCCCA  2026
Qy    554   ACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCCAGCTGGCTGCGCTGATCGGGCCCCA  613
Db    2027  ACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCCAGCTGGCTGCGCTGGGCCCCA  2086
Qy    614   CCCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC  673
Db    2087  CCCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC  2146
Qy    674   CCATCAGCCCCATCGAGACCGTGCCCGTGAGCTGAAGCCCCAGCATGGACGGCCCCAAGG  733
Db    2147  CCATCAGCCCCATCGAGACCGTGCCCGTGAGCTGAAGCCCGGCATGGACGGCCCCAAGG  2206
Qy    734   TGAAGCAGTGACCCCTGACCGAGGAGATCAAGGGCCCCATCTGCGAGGAGA  793
Db    2207  TGAAGCAGTGGCCCCTGACCGAGGAGATCAAGGGCCCCATCTGCGAGGAGA  2266
Qy    794   TGGGAGAAGAGGAGGAGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT  853
Db    2267  TGGGAGAAGAGGAGGAGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT  2326
Qy    854   TCGGCCATCAAGAAGAAGAAGACACCAAGTGGCACCAAGTCGGCTGGACTTCCGCGAGCTGA  913
Db    2327  TCGGCCATCAAGAAGAAGAAGACACCACCAAGTGCGCAAGTCGGCTGGACTTCCGCGAGCTGA  2386
Qy    914   ACAAGGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCGCCGCCCGGCCTGA  973
Db    2387  ACAAGGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCCCGGCCTGA  2446
Qy    974   AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGG  1033
Db    2447  AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGG  2506
Qy    1034  ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCGT  1093
Db    2507  ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCCG  2566
```

```
Qy  1094  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGAAGGCGCAGCCCCAGCATCTTCC  1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2567  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGAAGGCGCAGCCCCAGCATCTTCC  2626

Qy  1154  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCGCCGCAACCCCGAGATCGTGATCT  1213
          |||||||||||||||||||||||||||||||||||| | |||||||||||||||||||||
Db  2627  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCCGCAACCCCGAGATCGTGATCT  2686

Qy  1214  ACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGCGCAGCGACCGGCGCCCAAGATCG  1273
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2687  ACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGCGCAGCGACCGGCGCCCAAGATCG  2746

Qy  1274  AGGAGCTGCGCAAGCACCTGCGCTGCGGCTGGGGGCTTCACCACCCCCGACAGAAGCAGCACCAGA  1333
          |||||||||||||||||||||||||||||||          |||||||||||||||||||||
Db  2747  AGGAGCTGCGCAAGCACCTGCGCTGCGGCTTCACCACCCCCGACAGAAGCAGCACCAGA  2806

Qy  1334  AGGAGCCCCCTTCGTGGATGGGCTACGAGCTGACAGCTGCACCCGACAAGTGACCGTGCAGC  1393
          |||||||||||||||||         |||||||||||||||||||||||||||||
Db  2807  AGGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCGACAAGTGACCGTGCAGC  2860

Qy  1394  CCATGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAGACTGCTGGTGGGCA  1453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2861  CCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAGACTGCTGGTGGGCA  2920

Qy  1454  AGCTGAACTGGGCCAGCCAGATCTCACCCCGGCCATCGTTGCGCCAGCTGCTGTGCAAGCTGC  1513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2921  AGCTGAACTGGGCCAGCCAGATCTCACCCCGGCCATCAAGGTGCGGCCAGCTGTGCAAGCTGC  2980

Qy  1514  TGCGCGGCGCCAAGGCCCTGACCGAGATCTGCGCGGCGGGCCGTGCCCTGACCGAGGCCGAGC  1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2981  TGCGCGGCGCCAAGGCCCTGACCGAGATCGTGCCCCTGACCGAGGCCGAGC  3040

Qy  1574  TGGCCGAGAACCGGCCCGAGATCCTGCGCGGAGCGCGCGTGTACTACGACCCCAGCA  1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3041  TGGCCGAGAACCGGCCCGAGATCCTGCGCGGAGCGCGCGTGTACTACGACCCCAGCA  3100

Qy  1634  AGGACCTGGTGCCGAGATCCAGAAGAGCCACGACCAGTGACCAGATCTACC  1693
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3101  AGGACCTGGTGCCGAGATCCAGAAGAGCCACGACCAGTGACCAGATCTACC  3160

Qy  1694  AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACCGCCACGATGCGCACCCGCCCCACCA  1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3161  AGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACCGCCACGATGCGCACCCGCCCCACCA  3220

Qy  1754  ACGACGTGAAGCAGCTGACCGAGGCCGTGCCAGAAGATCGCCATGGAGAGCATCGTGATCT  1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3221  ACGACGTGAAGCAGCTGACCGAGGCCGTGCCAGAAGATCGCCATGGAGAGCATCGTGATCT  3280

Qy  1814  GGGGCAAGCAACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAGACCTGGGACACCTGTGGA  1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3281  GGGGCAAGCAACCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAGACCTGGGACACCTGTGGA  3340

Qy  1874  CCGACTACTGGCCAGGCCACCTGAGTCCGAGTTCGTGAACACCCCCCCCCTGG  1933
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3341  CCGACTACTGGCCGAGGCCACCTGAGTCCGAGTTCGTGAACACCCCCCCCCTGG  3400

Qy  1934  TGAAGCTGCTGGTACCAGCTGGAAGAGCCCATCATCGGCCGCGAGACCTTCTACGTGG  1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3401  TGAAGCTGTGGTACCAGCTGGAAGAGCCCATCATCGGCCGCGAGACCTTCTACGTGG  3460

Qy  1994  ACGGCGCGCCAACCGCGCCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGGCGGGCC  2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3461  ACGGCGCGCCAACCGCGCCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGGCGGGCC  3520

Qy  2054  GGCAGAAGATCGTGAGCCTGACCGAGACCAACCAGAAGCCGAGCTGCAGGCCATCC  2113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3521  GGCAGAAGATCGTGAGCCTGACCGAGACCAACCAGAAGCCGAGCTGCAGGCCATCC  3580

Qy  2114  AGCTGGCCCTGCAGGACACAGCGCGCAGCGAGGTGAACATCGTGACCGAGCAGCCAGTACCCCC  2173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3581  AGCTGGCCCTGCAGGACACGGCGCAGCGGCAGCGAGGTGAACATCGTGACCGAGCAGCCAGTACCCCC  3640

Qy  2174  TGGGCATCATCCAGGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCG  2233
```

```
Db  3641  TGGGCATCATCCAGGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCG  3700
          ||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2234  AGCAGCTGATCAAGAAGGAGAGGTGACCTGAGCTGGGTGCCGCCCACAAGGGCATCG  2293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3701  AGCAGCTGATCAAGAAGGAGAGGTGACCTGAGCTGGGTGCCGCCCACAAGGGCATCG  3760

Qy  2294  GCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGG  2353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db  3761  GCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCCTGG  3820

Qy  2354  ACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCG  2413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3821  ACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCG  3880

Qy  2414  GCGGCCCTAGG  2424
          |||||||||||
Db  3881  GCGGCCCTAGG  3891

Search completed: April 10, 2004, 21:24:41
Job time : 602.333 secs
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: April 10, 2004, 06:34:42 ; Search time 4148.08 Seconds
                                          (without alignments)
                                          17774.420 Million cell updates/sec

Title: US-09-610-313-30
Perfect score: 2469
Sequence: 1 gtcgacgccaccatggccga........gggctagcaccggtgaattc 2469

Scoring table: IDENTITY NUC
              Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14991090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*
              1: em_estba:*
              2: em_esthum:*
              3: em_estin:*
              4: em_estmu:*
              5: em_estov:*
              6: em_estpl:*
              7: em_estro:*
              8: em_tc:*
              9: gb_est1:*
             10: gb_est2:*
             11: gb_htc:*
             12: gb_est3:*
             13: gb_est4:*
             14: gb_est5:*
             15: em_estfun:*
             16: em_estom:*
             17: em_gss_hum:*
             18: em_gss_inv:*
             19: em_gss_pln:*
             20: em_gss_vrt:*
             21: em_gss_fun:*
             22: em_gss_mam:*
             23: em_gss_mus:*
             24: em_gss_pro:*
             25: em_gss_rod:*
             26: em_gss_phg:*
             27: em_gss_vrl:*
             28: gb_gss1:*
             29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description		
1	90.8	3.7	2598	11	AY103647	AY103647 Zea mays		
2	85.6	3.5	951	12	BM321451	BM321451 rockefell		
c 3	85	3.4	869	14	CK159167	CK159167 FGAS04056		
c 4	80.2	3.2	1132	12	BM320864	BM320864 rockefell		

No.	Score	%	Length	DB	ID1	ID2	Description	
5	79.6	3.2	1165	12	BM320900	BM320900	rockefell	
6	79	3.2	867	12	BM321430	BM321430	rockefell	
7	76.6	3.1	1550	12	BM321022	BM321022	rockefell	
8	75.6	3.1	671	13	CA093222	CA093222	SCCCCL300	
9	75.4	3.1	2299	11	AY106831	AY106831	Zea mays	
10	75	3.0	862	12	BM321023	BM321023	rockefell	
11	74.6	3.0	914	28	BZ568300	BZ568300	pacs2-164	
12	74.4	3.0	3134	11	AY109500	AY109500	Zea mays	
c 13	73.8	3.0	889	14	CK159613	CK159613	FGAS04107	
c 14	73.6	3.0	757	29	CC678788	CC678788	OGWII71TV	
15	73.2	3.0	853	12	BM321393	BM321393	rockefell	
16	72.8	2.9	566	12	BM587428	BM587428	170006873	
c 17	72.6	2.9	788	14	CB643171	CB643171	OSJNEb03L	
18	72.4	2.9	766	14	CC675888	CC675888	OGWCO51TH	
c 19	72.4	2.9	801	12	BJ375394	BJ375394	170006873	
20	72.2	2.9	688	14	CB648640	CB648640	OSJNEb12C	
21	72.2	2.9	764	14	CB651670	CB651670	OSJNEb16L	
22	72.2	2.9	766	14	CB642928	CB642928	OSJNEb03F	
23	72.2	2.9	809	14	CB641397	CB641397	OSJNEb01A	
24	72.2	2.9	841	14	CB651502	CB651502	OSJNEb16H	
c 25	72.2	2.9	851	29	CG260623	CG260623	OGWGE95TV	
c 26	72.2	2.9	860	29	CC691066	CC691066	OGUDF69TV	
c 27	71.8	2.9	1933	11	AY325173	AY325173	Rattus no	
28	71.8	2.9	5622	29	AY400359	AY400359	Homo sapi	
29	71.2	2.9	646	12	BM645449	BM645449	170006873	
30	71.2	2.9	724	12	BM584191	BM584191	170006872	
31	71	2.9	725	14	CB630209	CB630209	OSIIEb07A	
32	71	2.9	745	14	CB641936	CB641936	OSNEb01N	
33	71	2.9	766	14	CB629976	CB629976	OSIIEb06K	
34	71	2.9	787	14	CB658685	CB658685	OSNEC14O	
35	71	2.9	792	14	CB618305	CB618305	OSIIEa02B	
36	71	2.9	810	14	CB618374	CB618374	OSIIEa02D	
37	71	2.9	823	14	CB627740	CB627740	OSIIEb02O	
38	71	2.9	865	14	CB628660	CB628660	OSIIEb04F	
c 39	71	2.9	925	29	CN50091P	AL053013	Drosophil	
40	70.6	2.9	703	14	CB675843	CB675843	OSJNEe11M	
41	70.6	2.9	710	14	CB676048	CB676048	OSJNEe12B	
42	70.6	2.9	712	14	CB647546	CB647546	OSJNEb10I	
43	70.6	2.9	714	14	CB651194	CB651194	OSJNEb16B	
44	70.6	2.9	715	14	CB646590	CB646590	OSJNEb09C	
45	70.6	2.9	715	14	CB669855	CB669855	OSJNEe02K	

ALIGNMENTS

RESULT 1
AY103647
LOCUS   AY103647   2598 bp   mRNA   linear   HTC 16-OCT-2002
DEFINITION Zea mays PCO142084 mRNA sequence.
ACCESSION   AY103647
VERSION   AY103647.1 GI:21206725
KEYWORDS   HTC.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 2598)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE   Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2 (bases 1 to 2598)
  AUTHORS   Coe,E.H.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT   If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the

```
        maize cDNA sequences is either Virginia Walbot, Stanford or Pat
        Schnable, Iowa State, then clones may be requested from ZmDB:
        www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source          1..2598
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:638378"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
ORIGIN

Query Match        3.7%;   Score 90.8;  DB 11;  Length 2598;
Best Local Similarity  43.3%;  Pred. No. 0.21;
Matches  719;  Conservative   0;  Mismatches  909;  Indels   33;  Gaps    5;

Qy    170 GCGGCAAGAAGGAGGGCCACCAGATGAAGGACTGCACCGAGGCCAGGCCAACTTCTTCCGCG  229
          |||   |||      |||  | ||||| | ||| |||| |  ||||| ||| | ||| ||
Db    351 GCAGGCAGTGAGATCCTCGACTGCATCGCCCACGGCGGCGACATCTACGGCGTCACCACCG  410

Qy    230 AGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACCGCGGCCA  289
          |||||  |  | |||  |||||  |  |||| ||| | ||| | |||||| |||  |
Db    411 GCTTCGGCGGCACCTTCCACCGGCGCCACCGGCCTGCAGGTCAGGTCGAGCTGC  470

Qy    290 ACAGCCCCACCAGCCGCGAGCTGCAGGTGCGCGGCACCGGGCGCACAACCCCCGCAGCCGCGCG  349
          ||  | | ||  || |||| ||| || |||  |||||  | |  | ||||  ||  ||||| ||
Db    471 TCAGGGATCTCAACGCGGAATCTTCGGCCACCCGGAATCTTCGGCACCTGCGTCCGTCG  530

Qy    350 CCGAGCGGCCAGGCAGCCACCTGTGCAGCGGCCCCCTGGTGA  409
          || |||  | |  | ||| ||||| | ||||  | ||| ||
Db    531 AGGTCACCGCGGCGGCGGATGCGTGGCCATCAAGAGGGCCCTGCGCATCAACACCGTACTCCGGCA  590

Qy    410 GCATCAAGGTGGGCGGCCGCCAGATCTTCGGAGGCTTCGAGGAGGTGTGCTCAACACCGTGTCTGCTGCTGCCGCCC  469
          |||||| || |||  |||||| || | ||||| || |
Db    591 TCCGCTTCGAGGATCCTTCGGAGGCCTGCCGGGCATCCAACGAAGGTGCTGCTGAGAGCTGCGGGCCCCTGCC  650

Qy    470 TGCTGGAGGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCG  529
          | ||||| ||  ||  |||  |||||  |||  |||  |||||  ||||| |
Db    651 TGCCGCGTTCCGCGGGGCACCATCATCACGCCTCCGGGTTCCCGCTCTCCTTACATCGCCG  710

Qy    530 GCTTCATCAAGGTGCGGCCAGTACGACGACCGCGCGACATCCTGATCGAGACGAGGAGGCCA  589
          || ||||||  | | | |  | ||    |  || || ||| |||||
Db    711 GCCTCATCATCACCGGGCCGCGCCCCACCGGCCGTCACCG---TCGACGGGAGAGGTGG  767

Qy    590 TCGGCACCGTGCTGATCGGGCCGTTCAAGATCGCCGGCAGCGGCGGTTCTTCAAGCTCAACCCA  649
          ||||| || | ||||  |||  ||  |  ||  |||  || ||| | ||||  |||  |||| | ||
Db    768 TGATCGAGAAGCTGCGGCGCCGGTGTGGTGACCACGCGGCCGGCGCGG---GCGACGCGGG  827

Qy    650 AGCTGGGCTGCAGCCACCTGCCATCGTCAAGATCTTCCCCATCAGCCGCACGGCACGCTCCGGCTGCCGCTCCGTGAAGTCCGTGCCCTGAGAAGTGAGCTGA  709
          ||||| |  ||  || |||||   |||    |||  || |   ||| |
Db    828 AGGAGGGCCTCGACCATCGTCAAGATCTTCCCCATCAGCCGCACGGCTCCGGCTGCCGCTCCGTG  887

Qy    710 AGCCCGGCATGACGGCCAAGCCCGCCCAAGTGAAGCAGTGCCCCTGACCGGAGGAGAAGATCAAGG  769
          || ||||| ||  |||  | || |||  || |||  |||  |||  ||||||||| |||||  || |
Db    888 TGTACGACGCCAACGTCTGCGGGCCGTGTTCGCCATCAAGAGCACATCCTGGATGGCAGCTGCTCCTTCATGAAGC  947

Qy    770 CCCTGACACGGCCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGAATCCGGCCCCG  829
          |||||| |  |||||| ||| ||||| |||  |   ||| | ||   ||   |    ||  | ||
Db    948 TCATGAACGGCAAGCCCGAGTACAACGGCCAGTACACCCTGCGCCACCACCAAGTCGAAGCACCACCCCG  1007

Qy    830 AGAACCCCTACAAACACCCCCGTGTTCGCCATCAAGAGAAGACACCAAGTGCGCA  889
          |||  ||||| |||  ||   |  || |  |   || | | ||| |||   |||
Db    890 AGCTGGTGGACTTCTCGGGAGAGTGCAGCAAGCAACAAGCGCACCCAGGACTTCTGGGAGCTGCAGCTGGA  949

Qy    890 GGTTCCATCGAGCGCCGCGCCATCATGGCCGCATCCTTGGATGGCAGCTGCAGCTGCTTCATGAAGC  949
          |||  |  |  || |||| |   | ||| |||  ||  |||  |   ||||
Db   1068 AGGCCAAGAAGTTGAACGAGCTGACCCGC----------TGCTGAAGCCCAAGCAGG  1115

Qy    950 GCATCCCCACCCCGGCCGCCTGAAGAAGAAGAAGAGAGCGGTGACCGTGCTGGACGTGGGGCG  1009
          || |||  || |  ||| | ||   | || | || || || || | ||||||  | ||||
Db   1116 ACAGGTACGGCGCTCCGCACGTGCGCCAGTGGCTGGGCCCCAGATCGAGGTCATCCGCG  1175

Qy   1010 ACGGCCTACTTCAGCGTGACGAGGACTTCGCAAGTACACCGCCTTCACCATCC  1069
          ||  |  |  || ||  |  |||  | |  | || | || |  ||| |  ||
Db   1176 CCGCCACCAAGTCCATCGAGCGCGAGGTCAACTCCGTGACGACAACCCGGTCATCGACG  1235

Qy   1070 CCAGCATCAACAACGAGACCCCGGCATCCGTACCAGTACAAGTGCTGCCCCAGGGCT  1129
          || | ||| ||||| | ||| | |   |  |  | ||||  || ||||  ||  ||
Db   1236 TCCACGCGGCAAGGCGGCTGCACGGCGGCAACTTCCAGGGCACCCCATCGGCGGTCCA  1295

Qy   1130 GGAAGGGCAGCCCAGCATCTTCCAGACGCAGCAGCATGACCATGAAGATCCTGGAGCCCTTCCGCG  1189
          |  ||| |||  ||| |||||| || |  ||||| || | |  |||||| || |  ||| || | |
Db   1296 TGGACAACGCCCGCCCTCGCCATCGGCAACATCGGCAAGTCGGCAAGCATCGGCAGTCATGTTCGCCAGTTCTCCTCCG  1355

Qy   1190 CCCGCAACCCGGATGTGATCTACACGATGCATGGACCTGTACGTGGGCAGCGA--  1247
          || | ||||  | |||  |||  |||| || | |  || |  || ||| || |
Db   1356 AGTCGTCAAGCAGTTCTACACAACAACGGGTTCACCCACCTGGCGGCGGCGCCGCAACC  1415

Qy   1248 ----CCTGGAGATCGGCCAGCACGCACCGGCGCCAGATGCGGCAAGATCGGCCATGGTCTCTACTGCTCCGAGC  1303
              ||| || |||||  || |  || |  || ||| | |  ||  || ||| | |  |||||
Db   1416 CCAGCCTGCTGGACTACGGGCTTCAAGGGCACACGCCGATGATCGCCTCTACTGCTCCGAGC  1475

Qy   1304 GCTGGGGCTTCACCACCCCGACAAGAAGCACCAGAGCACCAGAGGACCGCGCCA  1363
          || |  |  || || || ||| ||   ||||   | |  || | | ||  |
Db   1476 TCCAGTACCTGGCAACCCCATCACCAACCGTGCAGAGCAGCAGCAACCAGG  1535

Qy   1364 GCTACGAGCTGCACCCGCACAAGTGGACCGTGCAGCCCATGAGCTGCCGGAGAAGGAGA  1423
          |  || ||||||||  | | ||||||  | ||  | |   || |||  ||||   |
Db   1536 ACGTGAACTCCCTGGGCTCGCTCGTCCGCGGACAGAGACCGCAGGGCGATCGACATCCTGA  1595

Qy   1424 GCTGGACCGTGAACGACATCCAGAAGCTGGTGGCAAGCTGGTCGGCAGTCGCCAGATCT  1483
          | ||| || ||||||| ||| | || ||||| | |  |||  |  |  || ||| | |
Db   1596 AGCTCATGTCGTCCACCTACATGTCGGCCAGCTGTGCGCAGCGCTGGCCAGGCCGTCGCTGGACCTCG  1655

Qy   1484 ACCCCGGCATCAAGGTGCCGCAGCGTGTGCAAGCTGGTGGCGCGCGGCCCAAGGCCCTGACCG  1543
          || |||  | ||||  | || | || | ||  | || ||  |  || |||  ||| ||||||
Db   1656 AGGAGAACATCAAGAGCGTCGGTGAAGAACAACACCGTGACCGACCGTGGCCAAGAGAGGTGCTGA  1715

Qy   1544 --------CCGCG  1594
          |||  || |||| ||||||  ||  |||| ||||   |||| |||  |||  || ||
Db   1716 CCATGAACCCCTGCGGGCGGCCGAGCTCTTCCAGCCGTGGTGTACGACGCCAAGGAGAGCTGATCAGCG  1775

Qy   1595 AGATCCTGGAGCGCCCGTGTACGACGCCAAGAAGAGGCCA  1654
          || ||||| ||||  ||  ||| |||    || ||| ||
Db   1776 CCATCGACCCGGAGGCCGTGTTCACGTACGCCGGAGGACCGGCCAGCTCCGTGCCGGCGC  1835

Qy   1655 AGATCCAGAAGCAGGGCGGCAGCGGCAACGGCGCTGGGTGGTGGACCACGCCACCCAGA  1714
          || |  || ||  ||| | |  |  |  || ||| |  || ||  |  |||| | || |
Db   1836 TGATCCAGAAGCTGCGCGCCGGCGTGTGCTGGTGGCCACCGTGCACCAGCGCG--GCGACGCGGG  1892

Qy   1715 ACCTGAAGACCGGCAAGTACGCCAAGATGCAGGCCCACCACCAACGTGAAGCAGC  1774
          ||  ||||| | || ||| |  ||   | ||  |  |||| ||| |||| | |
Db   1893 AGCGGGAGCCCTCCGTGTTCTCCAAGATCACCAAGGTTCTGGAGGAGGAGCTCCGCGCGGGTGC  1952

Qy   1775 TGACCGAGCCGTGCAGAAGATCGCCATGGAGAAGCATCGTG  1815
          ||||| ||  ||| ||| || |  |   |||  ||
Db   1953 TGCCCCAGGAGGTGGAGGCCCGTGGCGTGGCGCGGCAG  1993

RESULT 2
BM321451
LOCUS       BM321451                951 bp    mRNA    linear   EST 03-JAN-2002
DEFINITION  rockefeller.0.1211 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION   BM321451
VERSION     BM321451.1  GI:18055857
KEYWORDS    EST.
SOURCE      Mastigamoeba balamuthi
```

```
Db    816 CGGCCAGGCGGCTGCCGCGTCATCGTGACGGAGATCGACCCCATCTGCGCGCTGCGAGGCGGTC 875
          ||||| ||||| | | |||  || |||||||||||||||||||||| ||| ||||||||| |||
Qy   2394 CATGGACGACCTGTACGTGGCAGCGGCCGGGCCCTAGGATCGA 2435
          ||| | |||  ||||| | ||  ||||| |||  ||||||||||
Db    876 GATGGCGCGGCTTCGAGGTCAACACGCTCGAGGCGGGGCTCGA 917
```

RESULT 3
CK159167/c
LOCUS   CK159167   869 bp   mRNA   linear   EST 05-DEC-2003
DEFINITION FGAS040564 Triticum aestivum FGAS: TaLt5 Triticum aestivum cDNA,
      mRNA sequence.
ACCESSION CK159167
VERSION   CK159167.1 GI:38985053
KEYWORDS   EST.
SOURCE   Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
      Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 869)
 AUTHORS Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
      Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
      Links,M.G., McCarthy,B.L., Monroy,A., Muzak,I., Nilson,D.,
      Penniket,C., Roach,J.L. and Sarhan,F.
 TITLE   Functional Genomics of Abiotic Stress In Wheat and Canola Crops
 JOURNAL Unpublished (2003)
COMMENT   Contact: Wm L Crosby
      Bioinformatics
      University of Saskatchewan, Department of Computer Science
      1C101 Engineering Building, 57 Campus Drive, Saskatoon,
      Saskatchewan, S7N 5A9, Canada
      Tel: 306 966 1769
      Fax: 306 966 2033
      Email: fgas_ests@cs.usask.ca
      This sequence is the direct result of the Base calling software
      Phred (default parameters). It is the raw base calls. To aid in the
      identification of the high quality insert the software Lucy
      (default parameters) has been run on this sequence. Lucy identified
      the region [128,636].
      Plate: TaLt537 row: N column: 23.
FEATURES   Location/Qualifiers
   source   1..869
      /organism="Triticum aestivum"
      /mol_type="mRNA"
      /cultivar="Wheat line PI 178383"
      /db_xref="taxon:4565"
      /clone_lib="Triticum aestivum FGAS: TaLt5"
      /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
      subtractive hybridization) cDNA library from genotype
      PI178383 cold hardened at 2 C for 21 days and 49 days
      (equal amount of cDNA pooled together before subtraction,
      tester) and subtracted against genotype Norstar cold
      hardened at 2 C for 1 day (24 H)(driver). Modified Smart
      cDNA (Clontech)priming and non-directional cloning"
      /lab_host="DH5 alpha"
ORIGIN

Query Match   3.4%; Score 85; DB 14; Length 869;
Best Local Similarity 45.4%; Pred. No. 0.85;
Matches 304; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

```
Qy    580 AAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGAACATCATCGGCGCGCAAC 639
Db    827 ATGGCGGCGGCCCGGCACAACGAGGATCACCACCAGGAGGAAGAACACCACCAC 768
Qy    640 ATGCTGACCCAGCTGGGCTGACACCTGAACTTCCCCATCGACCCCATCGAGGACCGTGCCC 699
Db    767 CACAACGACGACAACGGCGACAACGAGCGCCACCACCAACAACAACAACAACAAC 708
Qy    700 GTGAAGCTGAAGCCCATGAAGCCCGGCCATCGATGGCGGCCATCGTGATCGTGATCGTACCGAGGAG 759
```

ORGANISM Mastigamoeba balamuthi
      Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE 1 (bases 1 to 951)
 AUTHORS Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
      Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
      Philippe,H.
 TITLE   The analysis of 100 genes supports the grouping of three highly
      divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
 MEDLINE 21819461
 PUBMED   11830664
COMMENT   Contact: Muller Miklos
      Laboratory of Biochemical Parasitology
      The Rockefeller University
      1230 York Avenue, New York, NY 10021, USA
      Email: mmuller@rockvax.rockefeller.edu
      Insert Length: 951 Std Error: 0.00
      POLYA=No.
FEATURES   Location/Qualifiers
   source   1..951
      /organism="Mastigamoeba balamuthi"
      /mol_type="mRNA"
      /strain="ATCC 30984"
      /db_xref="taxon:108607"
      /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
      /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.5%; Score 85.6; DB 12; Length 951;
Best Local Similarity 46.08%; Pred. No. 0.73;
Matches 323; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

```
Qy   1734 CGCCAAGATGCGCACCGCCACACCCCACCAACGACGTGAACGACGTGACCGAGGCCGTGCAGAA 1793
Db    219 GGCCAAGATCCGCTGGTGCTCGTGCAACATCTTCTGCAGCGAGGACCACGCGGCCGCCGGC 278
Qy   1794 GATCGCCATGGAGGAGCATCGTGATCTGGGCAAGAGAGACCCCCCAAGTTCCGCCTGCCCATCCA 1853
Db    279 CATTCGGCCAGGCCGGGCCGTCTTTCGCGTTGGAAGGGCGAGAACTTCCAGGAGTACTG 338
Qy   1854 GAAGGAGGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTG 1913
Db    339 GGGAGTGCACCTGAAGGCCCTGCTTCTTTCGGCCCTACCAGGGCCCTCAGATCATCGTCGA 398
Qy   1914 GGAGTTGCTGAACACCCCCCCCTGGTGAAGCTGTGTGTACCAGCTGGAGGAGAGGAGCCAT 1973
Db    399 CGACGGCGGTGACCGACTCTGATGATCCACAAGGGGTTCGGCGGCGAGGACAACCCCCAA 458
Qy   1974 CATCGGCGCCGAGACCTTCTACGTGACGGCGCGCCGCCAACCGGCGACGAGACATCGGCAA 2033
Db    459 GCTGCTGGAGGACGAGGGGCCTCGAGGAGGCCTCGGCGCTGCGGAGGAGTCGGCCTCAAGCA 518
Qy   2034 GGCCGGCTACGTGACCGTGACGACGCGGGCGCCGCAGAAGATCGTGAGCCTGACCGAGGACCACCAA 2093
Db    519 GGTCCAGAAGGAGGAGCAGCCCGGCTTCTGCACACAAGATCTCCCCGAGATCCGGGTGTCAG 578
Qy   2094 CCAGAAGACCGAGCTGAGGCTGCGAGGGCCATCCAGCCTGGCCCTGCAGGACAGCCGCAGCGGAGTGAA 2153
Db    579 CGAGGAGACGACGACGCTGGCGTNATGAGGCTGGCATCATCCAGCCCGACGCCCAGAGCAA 638
Qy   2154 CATCGTGACGACAGCCGACACGGCCGTGGGCATCATCCAGCCTGGCGTCAACGAC--TCTNTCACCAAGAGCAAGAGTTCGACAACATCTA 2213
Db    639 GTTCCCGGCCGTCAACGTCAACGAC--TCTNTCACCAAGAGCAAGAGTTCGACAACATCTA 695
Qy   2214 GACCGAGTCTGGTGAACCAGGATCATCGAGCAGTGATCAAGGAGGAGAGGTTACCTGAG 2273
Db    696 CGGCTGCCCCCACTCGCTCATCGACGGCATCAAGCGCGGCATCGGCACGACGCGTGCGCGG 755
Qy   2274 CTGGGTGCCCGCCCACAAGGGCATCGGCGCAACGAGCGAGATCGGTGGTGAGCAA 2333
Db    756 CAAGGTCGCCGTCGTCGCGGGCTTACGGCGGCTACGAAGGGGCTGCGCCCAGTCGTGCG 815
Qy   2334 GGGCATCCGCAAGGTGCTGCTTCCTGGGACGCGCATGATGCGGCGCATCGTGATCGTGATTACCAGTA 2393
```

```
Db   707 ACCAACAACAACAACAACAACAACAACAACAACGACAACCACCACCACCACCACCACCACAAC 648
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   760 AAGATCAAGGCCCTGACCGCCATCTGCCGGAGAGTGAGAAGGAGGGCAAGATCACCAAG 819
         ||| ||||| ||  |||
Db   647 AACAACAACGACAACAACAACAACAACAACAACACAACAACACAGAAGAAGGACAACAAC 588
                   ||||||  |||||
Qy   820 ATCGGCCCCGAGAAACCCTACACACACACCCGTGTTCGCCATCAAGAAGAAGGACAGCACC 879
         || |||||
Db   587 AACAACAACAACAACAACAACAACAACGACAACAACAACGACAACAACAACAACAACAAC 528

Qy   880 AAGTGGCGCAAGCTGGTGAGCTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAG 939
Db   527 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 468

Qy   940 GTGCAGCTGGGCATCCCCACCCCCGCCGGCTGGAAGAGGAAGAGGCGTGACCGTGCTG 999
Db   467 AACAACAACAACAACAACGACAACGACGACGACGACGACGAC 408

Qy  1000 GACGTGGGCGACGGCCTACTTCAGCGTGCCCTGGACGGGACTTCCGCAAGACCACCGGCC 1059
Db   407 GACAACAACGACAACAACGACAACAACAACGACAACAACAACAACAACAAC 348

Qy  1060 TTCACCATCCCCAGCATCAACAACGACAACCCCGCCATCCGCTACCAGTACAACGTGCTG 1119
Db   347 GACCACAACGACACAACGACAACAACAACAACGACAACAACCACCACAAC 288

Qy  1120 CCCCAGGGCGTGGAAGGGCAGCCCCAGCAATCTTCCAGAGCAGCAGCAACAACAACAAC 1179
Db   287 AACAACAACAACAACGACGACCAACCCGAGATCCGATCCTGATCTTCCAGTACATGGAC 228

Qy  1180 CCCTTCCGCGCCCCGCAACCCCGAGATCCGATCTTCCAGTACATGGACGACTGTACGTG 1239
Db   227 AACAACAACAACAACGACGACGACGTCAACGACAACAACGACAACAACGACCACGACAAC 168

Qy  1240 GGCAGCGAC 1248
Db   167 AACGGCAAC 159
```

```
Qy   139 CGGCGCCCCGCAAGAAGGGCTGCTGAAGAGTGCGGCAAGGAGGGGCACCAGATGAAGGAC 198
         |||  ||||  ||  ||
Db    79 CGCCGCGGAGGGCAAGACGGACTACCGCGCGCCAACTGGTGATCGATCCAGGACAACAAG 138
Qy   199 TGCACCGAGCGGCCAGGCCAATTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCAAGGCC 258
                                       |||||| |||||||
Db   139 TACAACAGGCCCCAGTACCGCTTCGTC------GTCCGCTTCACCACAGCGCGGGACATCGTC 192
Qy   259 CGGCAGTTCCCCAGCAGACAACCGGGCCAACAGCCCCACCAGCCGCGAGCTGCCAGGTG 318
         |||||| |||||||
Db   193 TGCCAGATCGGCTTCACGCCCAAGATCGACGGCCACCACATCCTTCCGCCGCCGCTACTCGCAC 252
Qy   319 CGCGGGCACAACCCCGCAGGCGAGGCCGGCGGGCCGAGCGCCAAGGCGCACCCTGAACTTCCCC 378
Db   253 GAGTCACCCGCTTCGGCGTCAAGCTCGGCTACAGCGCGTGGGCGGCGCCCAGATCAAGGAG 438
Qy   379 CAGATCACCCTGTGCAGCGCGCCCCCGTGTGCTGAAGAACCTGTCTGGAGGAGATGAGCGTGCCCCGGCAAG 498
Db   313 GGGCTGGCTGCTGGCCCCGCCGGTGTCGCCGACAACACCGTGCTCTGGAGGAGATGAGCTCGGGCAGG 372
Qy   439 GCCCTGCTGGACACCGGGCCGCCGACAACACCGTGCTCTGGAGGAGATGAGCTCGCCCGGGCAAG 498
Db   373 GTCAAGAAGGTCAACGGCGAGGACTACAACGTCGAGAGAGCTCGACGGCGCGCCCGT 432
Qy   499 TGGAAGCCCAAGATGATCGGCGGCATCGGGCGCCTTCATCAAGGTGCGCCAGTACGACCAG 558
Db   433 TTCAAGGCCCTGCTCGACGTCGGCGTCGGTCCGACCTCGACTGGCACCGACCCGTGTTCGGCC 492
Qy   559 ATCCTGATCGAGATCTCGCGGCAAGGAGGCCATCGGCGAGGCCGGCGGCGTCCCC----CACAGCGAGACCCGCTTC 618
         |||||||  ||||||  |||  ||
Db   493 GCCCTCAAGGGCATGTGCAGCGGCGGCGTCAACGTCCCC----CACAGCGAGACCCGCTTC 549
Qy   619 GTGAACATCATCGGCGCAACATGCTGACCGAGCTGACCCAGTTCACCCTGAACTTCCCCATC 678
Db   550 GTCGGCTTCAACGGCGACAAGAAGAGCTCAACGCCGGGTCCTCCGCAAGTACATCTTC 609
Qy   679 AGCCCCATCGAGACGGCCTGTCCCGTGAAGCTGAACGCGGCCCATGGACGGCCCCAAGTGAAG 738
Db   610 GGCGGCCACGTCGCCGGGTACATGAGCTCCTCAAGGAGCAGGAGCACGCCGCCGCCTTCGAC 669
Qy   739 CAGTGGCCCCTGACCGAGGAGGAAGATCAAGGCCTGACCGCCATCTGCCGAGGACGATGGAG 798
Db   670 CGCCAGTTCTCCGCGCTACGCCAAGGAGGGGTCACCGCCGACCCGCCACATCGTCCGAGAAGATCTAC 729
Qy   799 AAGGAGGCAAGATCACCAAGATCGGCCCCGAGAACCCCGTGTTCGCC 858
Db   730 ACCGAGGGCCCACAAGCAGATCGCCGCGACCGACGACCTTCGTCCCCAAGGCGCGCCTCGAAG 789
Qy   859 ATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGAGCTTCCGCGGAGCTGAACAAG 918
Db   790 CCCGAGGGGCGCCAAGCCCAAGCACTGGGCCAAGCGCAAGGCGCAGGCTGACGTACCACGACGCCAAG 849
Qy   919 CGCACCCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 966
Db   850 AACCGGTCGCCCAGAAGAGAGGTCCGCCTGGCGTACCCCGACGGCCCCC 897
```

KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
ORGANISM   Mastigamoeba balamuthi
           Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE 1 (bases 1 to 1165)
AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
           Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
           Philippe,H.
TITLE   The analysis of 100 genes supports the grouping of three highly
           divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE   21819461
PUBMED   11830664
COMMENT   Contact: Muller Miklos
           Laboratory of Biochemical Parasitology
           The Rockefeller University
           1230 York Avenue, New York, NY 10021, USA
           Email: mmuller@rockvax.rockefeller.edu
           Insert Length: 1165   Std Error: 0.00
           POLYA=No.
FEATURES   Location/Qualifiers
  source   1..1165
           /organism="Mastigamoeba balamuthi"
           /mol_type="mRNA"
           /strain="ATCC 30984"
           /db_xref="taxon:108607"
           /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
           /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.2%;   Score 79.6;   DB 12;   Length 1165;
Best Local Similarity 45.1%;   Pred. No. 4;
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;

Qy   90 CATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATGCCCCGCAACTGTCCGCGCCCCCG 149
           ||| || ||| || ||||||||||| ||||| |||| | ||| |||||||||
Db   46 CGTCAAGAACAAGGCGTACTTCAAGGCGTTCGTCGCGCCCAGTTCCGTGCGCGCGAGGG 105

Qy 150 CAAGAAGGCGTCGTGGAAGTGCGGCAAGGAGAGGGCCGCCAGAGATGAAGGACTGCACCGAGCG 209
           ||||| ||| |||||| | ||||||||| || |||||
Db 106 CAAGACCGACTACCGCGCGCGCCAACTGGTGATCCAGGACAGAAGCAAGTACAACAGCCC 165

Qy 210 CCAGGCCAACTTCTTCCCGGAGGACCTGGCCTTCCCCCAGGGCAAGGCCCCGCGAGTTCCC 269
           | ||||| ||||| ||||| |||||| ||||| | || |||
Db 166 CAAGTACCGCTTCGTC-----GTCCGCTTCACCACAGGACAATGCTGTGCCAGATCGC 219

Qy 270 CAGCGAGCAGAACCCGCGCCAACAGCCCCACACAGCCGGGCGAGCTGCAGGTGCGGGCGACAA 329
           ||||| ||| ||||| | || ||||||||| ||| |||||||| || |||||||||||
Db 220 CTACGCCAGATCGACGGCGACGGGCGACCACATCCTCGCCCGCGCGGGCGCACCCTCACCCG 279

Qy 330 CCCCCGCAGCGAGCGCCGGCGCCCAGCGCGCAGGGCCACCGGGCACTTCCCCAGATCACCCT 389
           ||||| ||| ||||| ||| |||||||| |||||| |||||| ||| ||||||
Db 280 CTTCGGCGCTCAAGCTCGGCCTCGACCAACTACGGCCGCGGCGACCTGGCCTGCTGCT 339

Qy 390 GTGGCAGCGCCCCCCGTGGTGAGCTATCAAGGTGGGCGGCCAGATCAAGGAGGGCCTGCTGGA 449
           ||| ||||| || || ||||| ||||| || ||||| ||||||||||||||| ||||||
Db 340 GGCCCGCCGTGTCGATGACGCTGAAGAAGCTCAACCTCTACTCCAAGTACGACGAGGGTGTCAAGAGAGGT 399

Qy 450 CACCGGCCGCCGCAGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAA 509
           || ||| ||| || || |||| ||||| || ||||| ||| |||| ||| |||||
Db 400 CAACGGCGAGGACTACACACGTCGCGGGCCTGACCAACGGGGCCCGCGCCCGCGTTCAAGGGCCC-- 457

Qy 510 GATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGA 569
           | ||||||||||||| ||||||| ||||| ||| || |||||| ||||
Db 458 --TGCTCGACGTCGGCCTTGGTCGGCCACCTCGACTGGCCGGGCCGTGTTCGCCGCCTCAA 515

Qy 570 GATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCGTGAACATCAT 629
           || || ||||| |||| ||||||||||| ||||| || ||||| ||
Db 516 GGGCATGTGGACCGGCGGCGTCAACGTCCCCACAGCGAGACCCGCTTCGTCGGCTTCAA 575

Qy 630 CGGCGGCAACATGCGGAGTGCGGCCTGGGCTGGCGCGCACCCTGCACCCTGAACTTCCCCATCAGCCCCATCGA 689
           |||||| | ||||| ||| |||| || ||||| || |||||| ||||||||
Db 576 CGGCGACAAGAAGGACTCAAGCGCCGGCGTCCTCGCAGTACATCTTCGGCGGCCCACGT 635

RESULT 6
BM321430
LOCUS   BM321430   867 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library
           Mastigamoeba balamuthi cDNA similar to ribosomal protein S4, mRNA
           sequence.
ACCESSION BM321430
VERSION   BM321430.1   GI:18055836
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
ORGANISM   Mastigamoeba balamuthi
           Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE 1 (bases 1 to 867)
AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
           Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
           Philippe,H.
TITLE   The analysis of 100 genes supports the grouping of three highly
           divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE   21819461
PUBMED   11830664
COMMENT   Contact: Muller Miklos
           Laboratory of Biochemical Parasitology
           The Rockefeller University
           1230 York Avenue, New York, NY 10021, USA
           Email: mmuller@rockvax.rockefeller.edu
           Insert Length: 867   Std Error: 0.00
           POLYA=No.
FEATURES   Location/Qualifiers
  source   1..867
           /organism="Mastigamoeba balamuthi"
           /mol_type="mRNA"
           /strain="ATCC 30984"
           /db_xref="taxon:108607"
           /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
           /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.2%;   Score 79;   DB 12;   Length 867;
Best Local Similarity 45.7%;   Pred. No. 4.4;
Matches 314; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

Qy 690 GACCGTGCCCGTGAAGCTGAAGCTGCGGCATGGACACGGGCCGGCCGCCGCCTTGACCGCCAGTTCTC 749
           |||||||
Db 636 CGCCGGCGTACATGAAGCTCCTCAAGGAGCAGCTCCAGGAGGAGGATGGAGGAAGGAGGGCAA 695

Qy 750 GACCGAGGAGAATCAAGGCCTTGACCGCCATCTGCGACGGCCGCCATCTGCGAGGAGGAGGATGGAGGAAGGAGGGCAA 809
           |||||||| || || || ||||||||||||||||||
Db 696 GCGGCTACGCCAAGGAGGGTGTCACCGGCCGACATGCTCGAGAGATCTACACCGAGGCCCA 755

Qy 810 GATCACCAAGATCGGCCCCGAGACCCTCACACACCCCGTGTTCGGCCATCAAGAGAGAA 869
           ||||||| ||||
Db 756 ------CAAGCAGATCCGGCGACCCGACCTTCGTCGTCCAAGCGGCGCTCGAAGCCCGA 809

Qy 870 GGACAGCACCAAGTGGCGCAAGTCGTGTGACTTCCGCGAGCTGAACAGGCCACCCAGGA 929
           ||| ||||| ||| ||||| ||| |||||| | || || ||||||||||
Db 810 GGGCGGCCAAGCCCAAGCACTGGGCGCAAGCCAGGCTGACGTACCACCACGACGGCAAGAACCG 869

Qy 930 CTTCTGGGAGGTGCAGCTGGGCATCCCGCCGCCGCCTGAAGAAGAAGAGAGCGT 989
           ||| | ||||| | ||||||||||||
Db 870 CGTCGCCCAGAAGAGGTCCGCTTGGCTACCCCGACGCCCCAAGGAGCCAGTAAATTCC 929

Qy 990 GACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGACGAGG 1039
           ||||||| | || ||| ||||| |||||
Db 930 CGGTGCCCTGTACACCGCCGCCATCTCCGGTGCGGCGTTGCCGCTGCTGCTNGCGG 979

```
source          1..1550
                /organism="Mastigamoeba balamuthi"
                /mol_type="mRNA"
                /strain="ATCC 30984"
                /db_xref="taxon:108607"
                /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                /note=syn: Phreatamoeba balamuthi"

ORIGIN

Query Match        3.1%;  Score 76.6;  DB 12;  Length 1550;
Best Local Similarity   45.6%;  Pred. No. 9.5;
Matches  308;  Conservative    0;  Mismatches  364;  Indels    3;  Gaps    1;
```

```
Qy   1761 CGACGTGAAGCAGCTGACCGAGCCGTGCAGAAGATCGCCATGCAGGAGCATCGTGATCTG 1820
               |||| |||| || |||||  |  || | |||   ||||||| | || |||| |||||
Db      5 CGGCACGGAGGACGGAGACCACCAGGCCGCATCGCGCAGGACCTGGTCTCGGGTCTT   64

Qy   1821 GGGCAAGACCCCCCAAGTTCCGCCTGCCCATCCAGGAGTACTTGGAGGACCTGGTGGTGAC 1880
            ||||| ||| | ||||| |||||| |  |||||  || |||  |||| |||| || |||
Db     65 CGCCTGGAAGGGCGAGAACCTCCAGGAGTACCTGGAGCGCGTGCTG  124

Qy   1881 CGACTACTGCAGGCACCTGGATCCCGGATCCCGGATCATGTCGACGACGACGACCCCCTGGT 1940
             ||| || || || |   ||   |||| | ||||||| | || |  || ||  |||
Db    125 CGGCCCCTACCAGGGCCCCGAGATCATGTCAGATCATGCTGACGACGGTGACCGGCCTGATGAT  184

Qy   1941 GAAGCTGTGGTACCAGTGGAGAAGGAGGCCATCATCGGCGCCGAGACCTTCTACGTGGA 2000
             || || | | || || ||| |||||| ||||  |  || |  |||| | | ||
Db    185 CCACAAGGGGTTCGCGGCCGAGGTTCGCGGGTGGGAGGACGACGAGGGCCCGA  244

Qy   2001 CGGGCCGCGCAACCGCGAGACCAAGATGGCAAGAGCCGGCTACGTGACGGCGACGGGGCCG 2060
            || || |  |||| | || | ||| |  ||  |||| || |||| || || || |||
Db    245 GGAGTGCGCTGCCTCAACAACGTGCTCAAGCAGGTCCAGAAGAGACCGAGCTGCAGGCCATCCA  304

Qy   2061 GCAGAAGATCGTGAGCCTGAGCCTGCCGATCCGCGGTTGCAGCGGAGACCAGCTGGCGTGATGAG 2120
            || |||| ||| || || || |  | ||  ||||||||| || ||| ||| ||| || ||
Db    305 GCACAGATCCTCCCGGATCCGCGGTGTCAGCGGAGGAGCATCGTGACGGAGACATCGGCGACGACTCC  364

Qy   2121 GCTGCCCGTCCTGCAGGACAGCGGCCAGCGAGGTGAACATCGTGACCGAGGTCACCGCTGTGGTGATCATCGA 2180
           || || |  | || || ||| || ||| ||  |  ||| |||| |  ||||||
Db    365 GCTGTACCAGGCTGACCCGCGACCCGCAGCTGCTGTTCCCGGCCGTCAACGTCAACGACTC  424

Qy   2181 GGGGCATCATCCAGGCCCGACAGTTTGACAACATCTACGGGTGGGGTGCCCGCCCACAAGGGCATCGG 2240
             ||| |||| |  || |  |||  |  |||| | |  ||||
Db    425 TG--TCACCAAGAGCAAGTTGACAACATCTACGGGTGGGGTGCGGGGCTACGG  481

Qy   2241 GCAGCTGATCAAGAAGAGGAGAAGGTGTACCTGAGCTGGGGTGCCCGCCCAGAGGGCATCGG 2300
            || | || |  |||  || ||| || || || || || ||| ||| ||| ||| || ||
Db    482 CATCAAGCGGCGCACCGGACGTGATCGTCCGGGCGCAAGGTCGCGCCGTTCGTCGGGGCTACGG  541

Qy   2301 CGGCAACGAGCAGATCGACACAAGCTGGTGAGCAAGGCCATCCGCAAGGTGCGTGTTCCTGGA 2360
            || |  |||  ||| || || | || |  |   || |||| |  || ||||
Db    542 CGACGTGGGCAAGGGCTGCCGAGTCGGTGCCCGGGCTGCCGGGGTCTGCCGGGTCATCGTGAC 601

Qy   2361 CGGGCATCGATGGGCGGCATCGTGATCTACCAGTACAATGACGACCTGTACGTGGGCAGCGG 2420
           || | ||| ||| || ||||   || || |||
Db    602 GGAGATCGACCCCATCTGCGCGGTCGCAGGGGTCGATGCGCGGGGCGGCTTCGAGTCAACACGCT  661

Qy   2421 CGGGCCTAGGATCGA 2435
            |||||| ||||  | |
Db    662 CGAGGCGGGGCTCGA  676
```

```
RESULT 7
BM321022        1550 bp    mRNA    linear    EST 03-JAN-2002
LOCUS     BM321022
DEFINITION  rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION   BM321022
VERSION     BM321022.1  GI:18055428
KEYWORDS    EST.
SOURCE      Mastigamoeba balamuthi
            Mastigamoeba balamuthi
  ORGANISM  Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1  (bases 1 to 1550)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., and
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE     The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
   MEDLINE  21819461
    PUBMED  11830664
  COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 1550   Std Error: 0.00
            POLYA=No.
```

```
FEATURES             Location/Qualifiers
```

(Partial alignment continued, top of page:)

```
Db    191 GCGCTGACCGACCCGGTGGTGAGGTCACCTGATCGTGATGCTGATGCAGCGGCCTGATCAAGATCGACGGC 250
Qy    178 GAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAACTTCTTCCGGCAGGACCTG 237
Db    251 AAGGTCCGCACCGACCACGCACTTCCCCGCGGGGGCTTCATGGACGTGTCTCGATCGACAAG 310
Qy    238 GCCTTCCCCCAGGCAGGCAAGGCCCGCGGAGTTCCCCAGCAGACAGCGCCC 297
Db    311 ACCGACGAGCACTTCCGCCTGCTCTACGACACCAAGGACGGGCGCTTCCAGGCGGCACCGCATC 370
Qy    298 ACCAGCCGCAGACTGCAGGT---GGGCGGGGACAACCCCCGCAGCGCAGGCGAGGCCGGCCAGG 354
Db    371 AACTCGACCGAGGCCAAGTTCAAGGTCGGCGCAAGGTCGGGCCGTGCAGCTGGCAACAAG 430
Qy    355 CGCCAGGGGCACCCTGAACTTCCCCCAGATCACCTGGTGGCAGCTGCCCCTGGTGAGCATC 414
Db    431 GGCATCCCGTACCTGTGACCCAGCACGCGCGCCGCACGATCCGCTCACCCCGACATC 490
Qy    415 AAGGTGGGCGGGCAGATCAAGGAGGCCCTGCTGGACACCGGCGCGCGGACACCGTGCTG 474
Db    491 AAGGTCAAGCAACGACACGGTCAAGATCGACCTGCGCTCGGGCAAGATCATCATGCGGCATCGGCCGGCTTC 550
Qy    475 GAGGAGGATGAGCTGCCCGGCAAGATGATCGGACACCGCAACCTTGCGCGTCGGGCGTC 534
Db    551 TTCGGAGATCGGCAACCTCGTCATGATCACTCGGCGACGCGCAACCTTGCGCGTCGGGGCATCGGC 610
Qy    535 ATCAAGGTGCGCGCAGTACGACCAGATCCTGATCGAGAGTCTCGCGGCAAGAAGGCCATCGGC 594
Db    611 ATTGTGCGGCCGGAGAGCACCAGGGCGTCGTTCGAGATCATCCACCTCAGGGACCACCCGTC 670
Qy    595 ACCGTCTGATCGGCCCCACCCCGTGAACATCATCGGCCGGCAACATGTCGACCCAGCTG 654
Db    671 GGCCACCAGTTGCGACGCGGGCCTGACCAACGTCTTCGTGATCGGCCAAGGGCACCAAGTCC 730
Qy    655 GGCTGCACCCTGAACTTCCCCATCAGCAGCCCCATCGAGACCGTGCCCGTGAAGCCC 714
Db    731 CTCGTCACGCTGCCCGCCGGCATCAAGAGAGTCGATCATCGAGGAGTTCCAGGCG 790
Qy    715 GGCATGGACGACGCCCCAAGGTGAAGCAG 741
Db    791 CGCCACCGCGCCAGGACCCAGAGGACCGAGCGAGGAG 817
```

```
RESULT 8
CA093222       671 bp    mRNA    linear    EST 23-SEP-2003
LOCUS     CA093222
DEFINITION  SCCCL3001F12.g CL3 Saccharum officinarum cDNA clone SCCCL3001F12
            5', mRNA sequence.
ACCESSION   CA093222
VERSION     CA093222.1  GI:34946529
KEYWORDS    EST.
SOURCE      Saccharum officinarum
            Saccharum officinarum
  ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

REFERENCE   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
             clade; Panicoideae; Andropogoneae; Saccharum.
             1 (bases 1 to 671)
AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE   The libraries that made SUCEST
JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT   Contact: Arruda P
             Centro de Biologia Molecular e Engenharia Genetica
             Universidade Estadual de Campinas
             Caixa Postal 6010, 13083-970, Campinas SP, Brazil
             Tel: 55 19 3788 1137
             Fax: 55 19 3788 1089
             Email: parruda@unicamp.br
             Clone distribution: clone distribution information can be found
             through the Brazilian Clone Collection Center (BCCC) at
             http://www.bcccenter.fcav.unesp.br
             Plate: 001 row: F column: 12
             Seq primer: M13/Reverse primer.
FEATURES   Location/Qualifiers
     source 1..671
             /organism="Saccharum officinarum"
             /mol_type="mRNA"
             /db_xref="taxon:4547"
             /clone="SCCCL3001F12"
             /lab_host="XL1Blue MRF'"
             /clone_lib="CL3"
             /note="Organ: Pool of sugarcane calli submitted to low
             (4oC) and high (37 C) temperature stress; Vector:
             pBluescript; Site_1: EcoRI; Site_2: XhoI; An
             unidirectional cDNA library generated from [Pool of
             sugarcane calli submitted to low (4oC) and high (37 C)
             temperature stress]. cDNA was prepared from polyA+ mRNA
             using ZAP - cDNA Synthesis Kit (Stratagene). The
             double-strand cDNAs were fractionated in a sepharose CL-2B
             40cm-columns and fragments sizing between 0.8 and 1.5 Kb
             were directionally cloned into the vector. Details of
             each source of RNA and library construction can be
             obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match   3.1%;   Score 75.6;   DB 13;   Length 671;
Best Local Similarity   47.3%;   Pred. No. 11;
Matches 228;   Conservative   0;   Mismatches 254;   Indels   0;   Gaps   0;

Qy   986 GCGTGACCGTGCTGACGTGGGCGACGCGCTACTTCAGCGTGCCCCTGGACGAGGACTTCC 1045
              ||| || |||||| ||| || || ||| || || ||||| |||||| || ||||| ||
Db   125 GCCTCAGCGAGGTCTACGGCGCCGCCGGCATTCGGCGGGGCCCAATGCCGGTGACTACG 184

Qy   1046 GCAAGTACACCGGCCGCCTTCACCATCCCAGCATCAACAACGAGACCCCCGGCATCCGCCTACC 1105
              |||||||| ||||| || || ||| ||||| ||| ||| ||||||||||||||||||
Db   185 GCGGCGAGATGGGGTTCCTTGACGTCGTCGAGCCCAAGCGGCGCGGCGCTCGATGACG 244

Qy   1106 AGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGA 1165
              |||| || |||| ||| ||||||||||| |||||| ||||||||||||||||||||||
Db   245 GCGTCGCCGCGGCCCTTCCGGCGCCTGGGGGTCTTGCAAGGTGGGAGCGGCGGCGGGTG 304

Qy   1166 CCAAGATCCTGGAGCCCTTCCGGCGCCGCAACCCCGAGATCGTGATCTACCAGTACATGG 1225
              | |||||||||| || ||||||| || || ||||| || | |||| || |||||
Db   305 CGTTTGGCCCTGGACCCGTCGTGCGGCGCAGCGGGCCGGGCCGTCCAAGAAGAGAGGG 364

Qy   1226 ACGACCTGTACGTGGCGCAGCGACCTTGGAGATCGGCAGCGCCAAGATCGAGGAGC 1285
              || || ||||| ||| || || |||||||| ||||| || || ||||||| ||
Db   365 TCGAGGGGATGGCGTCCAGAACCTTGATGCCGAGCCCGCGCCCGAGGCGCCTCAACG 424

Qy   1286 TGGCAAGCACCTGCTGCGTGGCGGGCTTCACCACCCCGACACCAAGCACCAGAAGGAGC 1345
              ||||| || || || |||||||| ||||| || |||||||| || |||| ||||||
Db   425 ACCGGCCTCTCCATGCGCGGCCCGTCGGTGCCGCGGCCAAGATCAGCAAGATGGGACGGG 484

Qy   1346 CCCCCTTCCTGGATGGGTCGGATGTCGTACGAGCTGCACGTCGACCTCGCAGCCCATCG 1405
              ||| || || || || |||||| ||| || ||| |||| ||| ||||| |||||||
Db   485 TCCTGGGCGACCGATCGACTACATGAAGGAGCTGCTGGAGAGGATCAAGTCTGCTGCATG 544

Qy   1406 AGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCATCCGAAGCTGA 1465
              ||||| ||||| || || ||| || || || || |||||||||||||||||| |||
Db   545 AGGAGATCGACCACCAGCAGCAGCAGGAGGAGGAGGGGCCGGCCGATGCCATGCATGCTCAACGTCTTCCGGGGCTCA 604

Db   1466 AC 1467
          ||
Qy   605 AC 606

RESULT 9
AY106831
LOCUS   AY106831   2299 bp   mRNA   linear   HTC 16-OCT-2002
DEFINITION Zea mays PCO142079 mRNA sequence.
ACCESSION   AY106831
VERSION   AY106831.1 GI:21209909
KEYWORDS   HTC.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 2299)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE   Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2 (bases 1 to 2299)
  AUTHORS   Coe,E.H.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT   If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES   Location/Qualifiers
     source 1..2299
             /organism="Zea mays"
             /mol_type="mRNA"
             /db_xref="MaizeDB:638374"
             /db_xref="taxon:4577"
             /clone_lib="Maize Mapping Project/DuPont Cornsensus
             Library"
             /note="this sequence is part of a project of EST
             assemblies resulting from the application of public
             contigs to seed DuPont contigs; this resource was
             assembled by DuPont as part of a collaboration for the
             overgo addressing of BACs in conjunction with the Maize
             Mapping Project"

ORIGIN

Query Match   3.1%;   Score 75.4;   DB 11;   Length 2299;
Best Local Similarity   43.0%;   Pred. No. 14;
Matches 734;   Conservative   0;   Mismatches 936;   Indels   37;   Gaps   6;

Qy   321 CGGCGACAACCCCGCCAGCGAGGCCGGCGAGCGCCAGGACCGCGAGCGCCCTGAACTTCCCCCA 380
              ||| ||| ||||||| || ||| ||||||||| || |||||||| ||| || |||||||
Db   256 CGGGTCCGACGGCCACACGTCCGTCGGGCCCTCGTGGCGTGCGGAGGTTCGCGTGCGGCCATGGCTCGTCCGCGAT 315

Qy   381 GATCACCCTGTGGCAGGCCCCGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGC 440
              ||||||||||| || || || || |||||||| || || || || |||||||| ||
Db   316 CAACACCCTCCTCCAGGGCTACTCCGGCATCCGCTTCGAGATCCTCGAGGGCCATCACCAA 375

Qy   441 CCTGCTGGACACCGGCCGCGGCGACGACACCGTGCTGGAGGAGGATGAGCGTCCGGGCAAGTG 500
              ||||||||| ||||| || || || |||||| || || || |||| ||||||| || ||
Db   376 GCTCATCAACACCGGCGGCCGGTCAGCCGTGCC-TGCCGCTGCGGCGCCACCATCACCGGTCGGG 434

Qy   501 GAAGCCCAAGATGATCGGCGGGCATCGGCGGGCTTCATCAAGTGCGCCAGTACGACCAGAT 560
              ||| || ||||| || ||| || || || ||| ||||||| |||| || || ||||||
Db   435 CGACCTCGTCCCGCTGTCCTACATCGGCGGGCCTCATCACGCGGGCGGCCTCCCAAGCGCCAGGC 494

```
Qy   561 CCTGATCGAGATCGCGCCAAGAGGCCATCGGCACCCGTGCTGATCGGCCGCACCCCCGT 620
             ||| |||||||| ||||||||
Db   495 C---GTCACGGTTGACGGCAGGAAGGTGACGCCCGCCGAGGCGTTCAAGGCGGCCGGCAT 551

Qy   621 GAACATCATCGGCCGCAACATGTCGGCCTTCTTCAAGCTCAACCCAGTCTGGGCTGCACCTGAACTTCCCCATCGA 680
             ||||| |||||| ||||||
Db   552 CGAGGGCGGCTTCTCCTGAAGCTGAAGCCCGGCATGGACGGCCCACGTCCTCGCCGTCGTC 611

Qy   681 CCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCCATGTGTCTTCGACGGCCACGTCCTCGCCGTCCTGTC 740
             |||||| |||||||||||||
Db   612 CGTGGGGTCGGCGCTCGCCATGAGTGCTTCGACGGTGCTTCGACGACCCAACGTCCTCGCCGTCCTGTC 671

Qy   741 GTGGCCCCTGACCGGAGGAAGATCAAGGCCCTGACCGCCATCGTCGAGGAGATGGAGAA 800
             ||||||| ||||||||
Db   672 CGAGGTCCGTGTCCGCGTCTTCTGCGAGTGATGAATGAACGGCAAGCCCGAGTACACGACGCCACCA 731

Qy   801 GGAGGGCAAGATCACCAAGATCGGCCCGAGAACCCCTACACACCACCCCGTGTTCGCCAT 860
             |||||||| ||||||
Db   732 CCTTACCCCACAAGTCAAGGCACCACCCCGGGG---------TCCATCGAGCGCGGCGGC 779

Qy   861 CAAGAAGAGGACGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCG 920
             |||  |||| ||||||||
Db   780 CATCATGATGAGCACATCCTGACTGCAGTCTCCTTCATGAAGCACGCAAGGAGTGAACGC 839

Qy   921 CACCCAGGACTTCTGGGAGTGCAGCTGGGCATCCCCACCCGCCGGCTCCGCAGCGCACGTG 980
             |||||||
Db   840 CATGGACCCGCTGCTGAAGCCGAAGCCGAAGCAGGCAGTACGGCTGTACCTTCAGGCTGCCCTGACGACGAGGA 899

Qy   981 GAAGAGCGGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGGCTGCCCTGACGACGAGGA 1040
             ||||||||||||
Db   900 GCTCGGCCCCCAGATCGAGTCATCCGCGCCCACCAAGTCCATCGAGCGAGAGGTCAA 959

Qy  1041 CTTCCGCAAGTACACCGCCTTCAACATCCCACGACATCAACACGAGACCCCGGCATCCG 1100
             |||||
Db   960 CTCGGTCAACGACAACACCCTGTCATCGACGCTCCACCGTGCAAGGCGTGCACGGCGGCAA 1019

Qy  1101 CTACCAGTACAACGTCGCTGCCCCAGGGCTGGAAGGGGCAGCGCCCAGCATCTTCCAGAGCAG 1160
             | |||
Db  1020 CTTCCAGGGGCACCCCGATCGGCGTGTCATGGCGTGTTCCATGGCGAACAGCGCCGCCTCGGCCAACAT 1079

Qy  1161 CATGACCAAGATCGTCAAGCCACTTGCTGCGGGCCTTCAACACCCCGAGGATGCGATCTACCAGTA 1220
             |||| ||
Db  1080 CGGCAAGTCATGTTCGCCCAGTTCTCCGAGCTGGTCAACGAGTTCTACAACAACCGGGGCT 1139

Qy  1221 CATGCGAGACCGTGACGTGGGCACGGA------CCTGGAGATCGGGCCAGCACCGCGCCAA 1274
             ||
Db  1140 GACCTCCAACCTGGCGGGCCAACCCGAGCCTGGACTACGGGTTCAAGGGCACCGA 1199

Qy  1275 GATCGAGGAGCTGCGCAAGCACTTGCTGCGCTGGGCTTCACCACCCCGACAAGAAGAGCA 1334
             |||||||
Db  1200 GATCGCCATGGGCAGGTCCTACTGGTCTCCGAGCTTCCAGTACCCTGGCCAACCCCATCACCAACCA 1259

Qy  1335 --------CCAGAGGAGCCCCCTTCGTGTGGCTGCGTAGCGGCTGCACCCCCGACAA 1385
             ||||
Db  1260 CGTCCAGAGCGCCGGAGCAGCAGAGCAGCAGGAACGTGAACTCCTCCTGGCGGCCTCGTCTCCGCCAG 1319

Qy  1386 GTGGTGACCGGTGCAGCCATCGAGCTGCCGGAAGGAGAGCTGCGGACCGTGAACGACATCCA 1445
             |||||||
Db  1320 GAAGCCGCCGCAGGGGGTGCAGACGTCCAGGGGGTGCTGACATCCTCAAGCTCATGTCGTCCACGTACATGTCGCGCT 1379

Qy  1446 GAAGCTGGTGGAGTGGGCCAAGCTGGTCGAACTGGGCACCTGGCCAGATCTACCCCGGGATCACACCCCGACAGTTCAAGGTGCCGCCA 1505
             |||
Db  1380 GTGCCAGGCCGTGAACCTGCCCCACCTGCGGGCGCCACTCGAGGGAACCTCAAGAGCGCCGTCAAGAGCTG 1439

Qy  1506 GCTGTGCAAGCTGCTGCTGCGGCCCTGCCCTGCCCCTGACCGGAGA 1565
             | |||||
Db  1440 CGTGATGGCCGGTGGCGTGGTGCTGACTGGACATCCTCAAGCTTCTGGCGGCGCACCTTCCACAGCGC 1499

Qy  1566 GGCCGAGCTGGAGCTGGACCGAGAACCGCGAGATCCTGCTGCGCGAGCGAGGCGCCGTGTA 1625
             |||||
Db  1500 GCGCTTCAGCGAGGCGAAGGCCCTGCTCACCGACCTCTCACCGCCATCGCCGCCAACCGCCGAGGCGCCGTCGACGTACTA 1559
```

```
Qy  1626 CTACGACCCCAGCAAGGACCTTGGTGGCCGAGATCCAGAAGCATCCGATCGAGACCAGAAGCAGGGCCACGACCAGTGAC 1685
             ||||||||||||
Db  1560 CGACGACCCCTGCAGCGCACCACTCGCCCCCTCAAGAACCTGAAGACCTGAAGAGAAGAGATCCGGGCCGTGCTGGTGA 1619

Qy  1686 CTACCAGATCTACCAGGAGCGAGCGAGCCCTTCAAGAACCTGAAGAACCTGAAGTACGCCAAGATGCG 1745
             |||
Db  1620 CCA------CGCCTCGCCAGCGGCGGAGGCCGAGGCGGAGGACGCCCAGCGCGTCCGTGTTCTC 1673

Qy  1746 CACCGCCCACCAACCACGACGTGAAGCAGCTGAAGCAGCTGCGCGAGGCGCTGCCCAGGAGATCGCCATGGA 1805
             ||||  |||
Db  1674 CAAGATCAACAGGTTCGAGGAGGAGTGCGGCGAGGCGGCTGCCCCAGGAGATGGGAGGCGCGC 1733

Qy  1806 GAGCATCGTGATCTGGGCAAGACACCCCAAGTTCCGCCTGCCCTGCCCATCCAGAGGAGGAGGACCTG 1865
             ||||||
Db  1734 CCGCGTGGCCTTGACACCGCCGATCAAGGATCAACAGGATCAAGGAGAGAGCCGGTC 1793

Qy  1866 GGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGGGCCACCTGGTGTCGTGAA 1925
             |||
Db  1794 GTACCCGCTTCACGCCGTTCATCCGCCAGGACCTCGGCGCCGTGTACCTGACCGGCGAGAA 1853

Qy  1926 CACCCCCCCCTGGTGAAGCTGTGTACCAGCTGAGAAGGAGGAGCCCATCATCGGCGCCGA 1985
             |||
Db  1854 GCTCAAGCTCCCCGGGAGGAGTGCAACAAGGTGTTCCTGCCCTCAGCGAGGGCAAGCT 1913

Qy  1986 GACCTTCTACGTGACGGCGCGCCAA 2012
             ||||
Db  1914 CATCGACCCCATGCTTCGACTGCCTCAA 1940
```

RESULT 10
BM321023

LOCUS   BM321023   862 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.594 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION   BM321023
VERSION   BM321023.1   GI:18055429
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 862)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 862   Std Error: 0.00
            POLYA=No.
FEATURES   Location/Qualifiers
     source   1..862
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.0%; Score 75; DB 12; Length 862;
Best Local Similarity 46.2%; Pred. No. 13;
Matches 283; Conservative 0; Mismatches 327; Indels 3; Gaps 1;

```
Qy  1734 CGCCAAGATGCGGCACCCGCCACCAACGAGCGTGAAGCAGCTGACCGAGGCCGTGCAGAA 1793
             |||| ||||||||
```

```
                                                                              /strain="2-164"
                                                                              /db_xref="taxon:287"
                                                                              /clone_lib="pacs2-164_7463"
                                                                              /clone_lib="pacs2-164"
                                                                              /note="clinical isolate 2-164 Whole genomic shotgun
                                                                                 library."

                               ORIGIN

                               Query Match       3.0%;     Score 74.6;   DB 28;   Length 914;
                               Best Local Similarity  46.8%;   Pred. No. 15;
                               Matches  270;   Conservative   0;   Mismatches  304;   Indels   3;   Gaps   1;

Qy    652  CTGGGCTGCACCTGAACTTCCCATCAGCCCATCGAGACGTGCCCGTGAAGCTGAAG  711
                |||||||||| | |   |||||  | |||||  ||  ||| | ||||   ||||||
Db    112  CTGCGCAACCAGCAGGACGGCCTGATCTTCAGCCAGCCGCCCTCGAGGTGCAGGCC  171

Qy    712  CCCGGCATGACGGCCCCAAGGTGAAGCAGTGGCCCTCCAGGAGAAGATCAAGGCC  771
                |||||| | ||| | |  |||||  | |  | | ||  || |||| | ||
Db    172  GGCAGCCTGGACAACCGCCAGCCGGCGCCAGGCGCCGACAACCAGAGTCGCGTATC  231

Qy    772  CTGACCGCCATCTGGCAGGAGATGAGGCCGCCGGTCAAGATCGGCCCCGAG  831
                ||||  |    || || | | |||  | |  ||||  |||| |
Db    232  GGTGGCGCGTTGGACAACCAGGGCGCGCCGGTGACAACCAGAGCCGGTAACCTAGACCTG  291

Qy    832  AACCCCTACAACACCCCGTGTTGCCATCAAGAAGAAGGACAGCAGCCAGTGGGCCAAG  891
                ||||  ||||||    | |  ||  ||| | ||| || |||| || |||  |||
Db    292  CAGAGCGGCAGCCTCGACACACGGCGGGCGTGCTCAACAGCGCCAAGGGGTTGGCTG  351

Qy    892  CTGGTGACTTCCGGCAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGC  951
                || |  |||  |  | ||  | | | | | |  |||||  |  || | | |||
Db    352  AAGCTGGTCACCGGGCTGTTCGACAACAGCGCCGGCGTCACCCAGGCCGTGCGCTGGAA  411

Qy    952  ATCCCCACCCCGCCGCCCGGCCTGAAGAAGAAGAGAGGGTGACGTGCTGGGCGGAC  1011
                ||| | | | |  | | || | |||| | | | | | | || |  | ||  |||
Db    412  ATCCGGCGCGGCAAGGCGTCGCCAACCAGCAGGGC---CATGTCTCGGCGTCGGCGGC  468

Qy    1012  GCCTACTTCAGCGTGCCCCTGACGAGGACTTCGACAACCAGGGTGGCGGCCTTCACCATCCC  1071
                 | ||| |  | | | | | | || ||  | || | | | ||  | || || ||
Db    469  GACAACCGACATCGTACCGGGCATCGTCCCGGCCACCGATTCGACATACACCGCACGCGGC  528

Qy    1072  AGCATCAACAACGAGACCCCGGCCACCGGCCTACCAGTACAACGTGCTGCCCGGGCTG  1131
                | ||  |  | | ||  || || || | || |||||| ||| | | | || |||
Db    529  CTGCTCAGCCTCGACGGCGACGCCTTCCTCAACCAGGCCGCCGGCGCCGGCGGGCGGC  588

Qy    1132  AAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCC  1191
                || | |||   ||  |  | || ||||| | | |  | | ||| |  || || | ||
Db    589  AAGGTCGCGGCCGGGCGGGCCGCATCGAGCTTCAGCCTGCGGCCTGGCCACCGCTTCGCC  648

Qy    1192  CGCAACCCCGAGATCGTGATCTACCAGTACATGGACG  1228
                | ||   | || | | | | |  ||||| ||| |||
Db    649  AGTTGGAAAGCCAGAGCGAGCTGCACCTGCACGCCCCG  685
```

RESULT 12
AY109500
LOCUS   AY109500   3134 bp   mRNA   linear   HTC 17-OCT-2002
DEFINITION   Zea mays CL1506_1 mRNA sequence.
ACCESSION   AY109500
VERSION   AY109500.1   GI:21213244
KEYWORDS   HTC.
SOURCE   Zea mays
  ORGANISM   Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 3134)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
               Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE   Maize Mapping Project/DuPont Consensus Sequences for Design of
               Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2 (bases 1 to 3134)
  AUTHORS   Coe,E.H.

```
Db    220  CGCCAAGATCCGCTGGTGCTCGTGCCAACATCTTCTCGACGCGAGGACGCGGCCGCCGC  279
                ||||| | | || || | |  | | | |  |  |  | ||  | | |  |  ||  |
Qy   1794  GATCGCCATGGAGAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCA  1853

Db    280  CATCGCGCAGGCCGGGAGTCTCCGGTCTTCGCCTCGAAGGGCGAGAACCTTCCAGGAGTACTG  339
                |||| | |||   || ||| | |||    |  | ||||| |  || |
Qy   1854  GAAGGAGACCTGGAGACCTGGTGGACCGACTACTGGCGAGGCCACCTGGACTCCCGAGTG  1913

Db    340  GGAGTGCACTTNGAAGGCCCTGTGCTTCGGCTTCGGCCCCTACCAGGGACCTGCAGGACCTCTCGA  399
                |||| | | |  ||  | | | ||  ||  |  | | |||  |||| | | | ||
Qy   1914  GGAGTTCGTGACACCCCCCCCTGGTGAAGCTGTGTGTACCAGCTGTGAGAAGGAGCCCAT  1973

Db    400  CGACGCGGTGACGCGACTTCTAATGATCCACAAGGGGTTCGCGCGGCGGACAACCCCAA  459
                 | | |  | |  |  | ||| | ||  || || | || || | | || |||||
Qy   1974  CATCGGGCCCGAGACCTTCTACGGTGACAGGGCCGCCGCCGCCGAGACGATGCGGCAA  2033

Db    460  GCTGCTGGAGGACGAGCAGGCCCTCGAGGAGTGCCTCGCTCTCAACAACGTGCTCAAGCA  519
                || | ||||| | | ||  |  |  | ||   | || || || | |||| || |  |
Qy   2034  GGCCGGCTACGTGACCGACACCGGGCGCCGGCCGCAGAAGATCCTGAGCTGACCGACCACCAA  2093

Db    520  GGTTCCAGGAGGACGAGCAGCCGCCGGCGGTTCTGGCCATCCAGCTGGCCATCCAGGGC  579
                ||| |||| | | | || ||||  || | || | ||| |||  | || ||| | |
Qy   2094  CCAGAAGACCGAGCTGCAGGGCTGCAGGGGTCCTGCCCTGCAGGACGCAGCGGCCAGGTGAA  2153

Db    580  CGAGGACGACGACTGGCGTGATGAGGGTGTACCAGTGCACGCGGACGGCCAAGCTGCT  639
                | ||| | | || |  | | ||| |  | ||||  | | |||| | | || || |
Qy   2154  CATCGTGACCGACGACAGCCCAGTACGCCCTGGCATC---ATCCAGGCCCGACAACAGAG  2210

Db    640  GTTCCCGGCCGTCAACGTCAAGCTTGTCTGTCACCAAGAGACAAGCAAGTTTGACAACATCTACGG  699
                || || | | | ||| | |  | |  || | | ||||  | | || | || ||||
Qy   2211  CGAGAGGGCAGCTGGTGAACCAGATCATCGAGACGCTGATCATCAAGAAGAAGAAGGTGTACCT  2270

Db    700  CTGCGCGCCACTCGCTCATGCAGCGCATCAAGCGCGCAACAAGGGCATCGTGGTCCGGCGCAA  759
                || | | || | || | | || |  |  || |  |   | | ||| | ||  | || |||
Qy   2271  GAGCTGGTGGCTGCCCCGCCCCAAAGGGCATCGCGGCGCAACGAGCAGCAGATGCGAACGAAGCAGTGGTGAG  2330

Db    760  GGTCGCCTCGTCGTCGGCGACGTGGGCAACGTGGGCAAGGGCGTACGGCGACGTGGGGCAAGGTGGCAAGG  819
                ||| | | | | | | | || | ||| || | | |  ||  | | || | |
Qy   2331  CAAGGGCATCCGC  2343

Db    820  CCAGGGCTGCCGC  832
                ||||| | ||||
```

RESULT 11
BZ568300
LOCUS   BZ568300   914 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION   pacs2-164_7463.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
               pacs2-164_7463, genomic survey sequence.
ACCESSION   BZ568300
VERSION   BZ568300.1   GI:27201058
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
  ORGANISM   Pseudomonas aeruginosa
               Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
               Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 914)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
               Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE   Whole-Genome-Sequence variation among multiple isolates of
               Pseudomonas aeruginosa library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
               Genome Center
               University of Washington
               Box 352145, Seattle, WA 98105-2145, USA
               Tel: 2062216954
               Fax: 2066857244
               Email: craymond@u.washington.edu
               Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..914
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"

```
TITLE      Direct Submission
JOURNAL    Submitted (25-APR-2002) Maize Mapping Project, University of
           Missouri, Columbia, MO 65211, USA
COMMENT    If you are interested in getting corresponding physical clones,
           these are publicly available from ZmDB and may be found by BLAST
           searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
           www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
           maize cDNA sequences is either Virginia Walbot, Stanford or Pat
           Schnable, Iowa State, then clones may be requested from ZmDB:
           www.zmdb.iastate.edu.
FEATURES   Location/Qualifiers
     source 1..3134
            /organism="Zea mays"
            /mol_type="mRNA"
            /db_xref="MaizeDB:630364"
            /db_xref="taxon:4577"
            /clone_lib="Maize Mapping Project/DuPont Cornsensus
            Library"
            /note="this sequence is part of a project of EST
            assemblies resulting from the application of public
            contigs to seed DuPont contigs; this resource was
            assembled by DuPont as part of a collaboration for the
            overgo addressing of BACs in conjunction with the Maize
            Mapping Project"
ORIGIN

Query Match        3.0%;   Score 74.4;  DB 11;  Length 3134;
Best Local Similarity   40.6%;  Pred. No. 20;
Matches   963;  Conservative  0;  Mismatches 1347;  Indels  61;  Gaps   8;
```

```
Qy    115 GAGGGCCACATCGCCCGCAACTGCGCGCCCCGCGCCCCGCAAGAAGGGCTGCTGAAGTGCGGC   174
          |||||| | |||  | |  |||     |||        |||||||||  ||||    ||
Db    329 GAGAGCTGTACGGCTCGCACCTGCAGCTGGAGGTGGACCCCGGGTGCAGGACGCGCGGCG   388

Qy    175 AAGGAGGGCCACCAGAGATGAAAGACTGCGACCGCGATCAGCGGAAGCCCAACTTCTTCCGCGAGGAC   234
          ||| |||  || |  |||||    ||||  |  |  |    |||     ||   |||||
Db    389 ACGAAGGCACCCGCGGACGCGTGATCGAGCGCCCGGTGATCAGCGTGCTGGTGAGGGTGACGGGGA   448

Qy    235 CTGGCCTTCCCCCAGGGCAAGGCCCGGGGCCAGTTCCCCAGCGAGCAGAACCGCGCCAACAGC   294
          ||| |||    |     ||  |   |    |||   |  |||||| || | ||  |
Db    449 AGCACCCGCTCAACTGCGAGCCGCCGCTAGCGGCGCGGGTG-ATGGCACCACGGCGCTTCATCACC   507

Qy    295 CCCACCAGCGCGAGCTGCAGGTGCGCGGCGACAAGCCGCGAGGCGCGGCCGGAG   354
          ||| |||      | | |   |  |||    ||||  |||||||  ||
Db    508 CCGGCGCGGGTGCCACTACGTGCGCAACCACGGCGCCGTGCCGCNNNNNACTGGGGCGACG   567

Qy    355 CGCCAGGGCACCCTGAACTTCCCCGAGATCACCCTGTGTGCAGCGCCCCGTCTTCCGTGTGAGCATC   414
          |||||||||||    |||||  ||  |   ||      | |      ||||||||||
Db    568 TGGACCGTGGAGGTGACGGGCCTTCGTGAGGGTCGGCAGCAGGAGCTGAAGGCTGACGAGCTG   627

Qy    415 AAGGTGGGCGGCAGATCAAGGAGGCCCGGGAGATCCCCGTGAGGTCGCGCAAGGCCGCGC   474
          ||| |||||   ||   |||||||| |||||   |   || | | |||
Db    628 GCGCGGCGACTTCCCCGCCTGGAGATCCCCGTGACGCTGCTCCCGTCGGCGCGGCGCAACCGCGCGC   687

Qy    475 GAGGAGATGAGCCTGCCGGCCAAGCAGCCCGGGAAGCCCAAGATGATCGGCGGCGCTTC   534
          |||||| |||||||||  ||   || |||||||||||  || | || ||
Db    688 AAGGAGCAGCAGAACATGTCAGCAGACGTGGGCTTCAACTGGAGCCCAAGAAGGCCATCGGC   747

Qy    535 ATCAAGGTCGCCAGTAGCACCAGATCCTGATCGAGATCGAGATCGGC---TGCGGCATCGTG   594
          ||||||| || |  |  ||||  ||   ||   |  | |   ||   |  |||||||||||
Db    748 ACGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCGTGCTCCGCGC--TGCGGCATCGTG   804

Qy    595 ACCGTGCTGATCGGCCACCCCGTGAACAATCATCGGCGCGAACATGTCGACCCAGCTG   654
          || | | | |  |   || ||  |  |  ||| | |  | ||||
Db    805 CCCCGCAAGGGCGCGCCCTCAACGTGTTCGTTCGAGGGGCGCCCAGNNNNNNNNNNNNNNN   864

Qy    655 GGCTGCGACCCTGAACTTCCCCATCAGCCCATCGACACGTGCCGTGAAGTCGAAGCCC   714
          |||||| |  ||| ||||   |   |||| |||| | || |  |  ||    |||
Db    865 NNNNNNNNNNNGTACGGCCACCCAGCTCACGCGAGTGGGGCCCTGGACCCGTCGGGGGGAC   924

Qy    715 AAGCAGTGAACTGGAGCAGTGAAGCAGTCAGTGCCATCTCGTGATCTCGGGGCAAGGCCCTG   774
          || |||||  ||||    | | |      ||| | ||       || |||   |||||
Db    925 ATCATGCTCGCTACATGCAAGCAGCAGATCAGCAACGCCGGTCGCCGACGAGCTTCCGCCGGAGGCTTCCCCGTG   984
```

```
Qy    775 ACCGGCCATCGCGAGGATGGAGGATGGAGGGGCAAGATCACCAAGATCGGCCCCGAGAAC   834
          |||||
Db    985 CGGGTCATCATCCCCGGCTGCATCGGTGCCGCCATGGTCAAGTGGCTCAAGCGC---ATC   1041

Qy    835 CCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAGCAGCACCAAGTGGCGCAAGCTG   894
          ||| | |  ||          |||       ||  ||    ||    |  |||||
Db    1042 ATCGTCANNNNNGCCGAGTCCGACAATTACTACCATTCAAGGACAACACCGGCGTCCTGCCG   1101

Qy    895 GTGGACTTCCGCGACTGAACAAGCGCACCCAGCAGACTTCTGGGGAGGTGCAGCTGGGCATC   954
          ||||   ||  ||     |  |||| ||   || || || | ||    |   |    |
Db    1102 TCGCACGTCGACGCCGGCAGCCTCGCCAACGAAGAAGAAGGAGAGGCGTGGTGGTACAAGCCGGAGTACATC   1161

Qy    955 CCCCACCCCGGCTGAAGAAGAGAGAGGGCTGACCGTGTTGGACGTGGGCGACGCC   1014
          |||  | |||    |||||  ||  | ||| |   |   |    | |   |  ||||||
Db    1162 ATCAACGAGCTGAACATAAACTCGGTGATAACGACGCCGGGGCACGACGAGATCCTGCCC   1221

Qy    1015 TACTTCAGCGTGCCCCTGGACGAGGACTTCGCAAGTACACCGCCTTCACCATCCCCAGC   1074
          |||||||| |    |  ||   | || ||  |  ||| ||    | |
Db    1222 ATCACAGCATTACCACACGCGGCGGCTACACCATGAAAGGATACGCCTACTCCGGCGGC   1281

Qy    1075 ATCAACAAGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCAGGGCTGGAAG   1134
          ||| |  |  | |||    ||  | |   ||| |      |||| || ||
Db    1282 GGCAAGAAGGTGACGGGGTGGAGGTGACGTGACGCTGGAGCAGCAATGGCTGTGTGTGTGC   1341

Qy    1135 GGCAGCCCCAGCATCTTCAGAGACAGCAGCAAGATCCGAGCCAAGCTGACCAAGCTGCCCGC   1194
          |||| |  | | ||  |     ||| |   | |  ||||| || || |  ||  | ||   |
Db    1342 CACCTGCAGCACCCGGAGAAGCCCAACAAGTACGGCAGTACTGGTGCTGGTGTGCTTCTCGG   1401

Qy    1195 AACCCCGAGATCGTGATCTACCAGTACATCATGGACGACGCCTGTACGTGGCAGCGACCTGGAG   1254
          || |||||| ||   || | |  ||    |  | ||||||  ||||
Db    1402 TCCGTCGAGGTGGAGGTCCTCGACCTCGGCGGCGTGCAGGCGCCCGGGTGCAGGACGCGCGGCGCATG   1461

Qy    1255 ATCGGCCAGCACCGCGGCAAGATCGAGGAGCTGCGCAAGCACTTGCTGCGCTGCGCTGGGCGCTTC   1314
          |||  || |  | | |  |  |  ||  |||  |||
Db    1462 GACCAGTCGCTCAACACCCGAGAGCTCATATGGAACCTCATGGGGATGATGAAC   1521

Qy    1315 ACCACCCCCACAACAGAAGCACCAGAGAGCAGCCCCCCTTCCTGTGATGGGCTACGAGCTG   1374
          |||||   | |  || ||  |||  || |  | |  |  |||   ||||||| ||  |||
Db    1522 AACTGCTGGTTCAAGGTGAAGGTGAACGTGTGCCCGTCCGCGCACAAGGCCGCAGATCGGGCTG   1581

Qy    1375 CACCCGCAAGTGACCGTGCAGCCCCATCGAGC---------TGCCCGAGAAGGAG   1422
          |||| |||| |||  |  |||   | | |||          || | ||||  ||
Db    1582 GTGTTCGAGCACCCGGACGCAGCCCCGCAACCAGCCCGGGCGGGTGATGGGCGGCGAGAAG   1641

Qy    1423 AGCTGGACGACCGTGAACACATCCAGAGTACATCCAGATC   1482
          ||| |||| ||    |   ||   |  ||   | | |
Db    1642 CACNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCAAGCCAGCACGTCCAGCGCCGTTC   1701

Qy    1483 TACCCCGGATCAAGGTGCGCCAGCTGTGCAAGGTGCTGCCCTGACC   1542
          |||| |||| ||| | |  |||  |  || |    | || ||||
Db    1702 ATGAACACCACGACGTCGGCAAGCAGTTCACCATGTCCGAGGTGCGCGAACGACGCGTCG   1761

Qy    1543 GACATCGTGCCCGACCGAGGAGGCGAGCTGGAGCTG------CCGAGAACCGCGAG   1596
          ||| |||||||  | | |||   |  || ||   ||      |||   ||  |  ||
Db    1762 CAGGAGTCGGCGGTGGATCGCGTGCACGGCGCCGTCTACGACTGCACCAAGTTCCTCAAG   1821

Qy    1597 ATCCTGCGCAGGCCCGTGCACGCGGGTGTACTACGACCCCCAAGGACCTTGGTGCCGAG   1656
          || | ||||||   || ||||     || ||     | ||||   ||
Db    1822 GACCACCCCGGCGGCGGCGGCGACCAGCAGCTCATCATCAACGCCTACCAGATTACCAGGAGCCC   1881

Qy    1657 ATCCAGAAGCAGGGCCACGACGACCTGGACTTCACCAAGGAGCCC---------   1707
          || || |   |||| ||  |   |  ||| |||||
Db    1882 TTCGACGCCATCCACTCGCACAAGGACGCCCAAGGCGCTCCTCGACACTACCGCGTCGGCGAG   1941

Qy    1708 TTCAAGAACCTGAAGACCGGCAAGGTACGGCAGCCCAAGATGCACCAACGACGTG   1767
          |||   | |  |||     ||   ||   ||  |  ||   || |
Db    1942 CTCATCACCACGGGCACCGGCACCGAGTGGCTCACAGCTCCGACAACTCCCGTCCAGCCGGCTCCTG   2001

Qy    1768 AAGCAGTGACGAGGACGCCGTGCAGAGCATCGTGATCTGGGGCAAGGCCCTG   1827
          ||||| |||  ||| | |  |   || |||   || |  ||   ||||
Db    2002 TCGCACCTGGCCTCGCGCCCATCCGCCGGAGCCCGGCGAGGAGGNNNNNNNGGGCTCTCCAACCCGCGCGAT   2061
```

This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [123,668].
Plate: TaLt539 row: N column: 07.

FEATURES   Location/Qualifiers
     source   1..889
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Wheat line PI 178383"
                     /db_xref="taxon:4565"
                     /clone_lib="Triticum aestivum FGAS: TaLt5"
                     /note="organ: Crown; Vector: pGEM-T; SSH (suppression
                     subtractive hybridization) cDNA library from genotype
                     PI178383 cold hardened at 2 C for 21 days and 49 days
                     (equal amount of cDNA pooled together before subtraction,
                     tester) and subtracted against genotype Norstar cold
                     hardened at 2 C for 1 day (24 H) (driver). Modified Smart
                     cDNA (Clontech) priming and non-directional cloning"

ORIGIN

Query Match   3.0%; Score 73.8; DB 14; Length 889;
Best Local Similarity 47.4%; Pred. No. 18;
Matches 251; Conservative 0; Mismatches 277; Indels 1; Gaps 1;

Qy   753 CGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCCAAGAT 812
              ||||| ||||||||||||||||||||||||||| |||||| |||
Db   674 CGAGACAACACCACCAACAACAACAACAACAGCAACAACAGCAACAGCAACAACAACACAG 615

Qy   813 CACCAAGATCGGCCCAGACCCTACACACAACAACAACAACAACAGCAACAACAAGAAGAGGA 872
              ||||| || || |||||||
Db   614 CAGCAACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAACAGCAACAACAAGAACAA 555

Qy   873 CAGCACCAAGTGGCGCAAGCTGGTGGACTTCGCGAGCTGAACAAGCGCCACCCAGGACTTG 932
              ||||| |||
Db   554 CAACAACGACAACAACAACAACAAGACAACAACAGCAACAACAGCAACAACAACAACAACAA 495

Qy   933 CTGGGAGGTGCAGCTGGGCATCCCC-CACCCCGCGCCTGAAGAAGAAGAGAGCGTGA 991
              |||| |||
Db   494 CAACAACAACAACAACAACAAGAGACAACAACAACAACAACAACAACAACAACAACAACG 435

Qy   992 CCGTGTGGACGTGGGCAAGCAGCGACGCTACTTCAGCGTGCCCCTGGACGACAACAACGACAACAACAACA 1051
              || ||||||||| |||
Db   434 ACGACGACGACAACGACGACGACGACAACAACGAGAACCCCGGCGTCAACGACGACAACAACAACGACAACAACA 375

Qy   1052 ACACCGCTTCACCATCCCGAGCATCAACAACGAGACCCCGGCGTCCGCTACCAGTACA 1111
              |||||||||||||||||||| |||
Db   374 ACAACGACAACAACAACAACAACAACAATAACGACAACAACGACAACAACAACAACAACA 315

Qy   1112 ACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGACGCAGCAGCATGACCAAGA 1171
              | |||||||| |||
Db   314 ACAACAACAACAATGACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 255

Qy   1172 TCCTGGAGCCCTTCCGGCCGCCAACCCCGAGATGTGATCTACCAGTCATGACGACC 1231
              ||||||||||||||||||| |||
Db   254 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACGACGACAACG 195

Qy   1232 TGTACGTGGGCAGCGACCTGGAGATGTGCCGAGCACCGCGCCAAGATCGA 1280
              |||||||||||||||||||||
Db   194 ACGACAACGATAACACGACAACAACAACAACGACAACAACAACGGCAACAACACAA 146

RESULT 13
CK159613/c
LOCUS   CK159613   889 bp   mRNA   linear   EST 05-DEC-2003
DEFINITION FGAS041079 Triticum aestivum FGAS: TaLt5 Triticum aestivum cDNA,
            mRNA sequence.
ACCESSION   CK159613
VERSION   CK159613.1 GI:38985955
KEYWORDS   EST.
SOURCE   Triticum aestivum (bread wheat)
  ORGANISM Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 889)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE   Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2003
            Email: fgas_ests@cs.usask.ca

RESULT 14
CC678788/c
LOCUS   CC678788   757 bp   DNA   linear   GSS 19-JUN-2003
DEFINITION OGWII71TV_ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0595K21,
            genomic survey sequence.
ACCESSION   CC678788
VERSION   CC678788.1 GI:32083564
KEYWORDS   GSS.
SOURCE   Zea mays

ORGANISM　Zea mays
　　　　　　Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
　　　　　　Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
　　　　　　clade; Panicoideae; Andropogoneae; Zea.
REFERENCE　1　(bases 1 to 757)
AUTHORS　　Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
　　　　　　Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
　　　　　　Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE　　　Consortium for Maize Genomics
JOURNAL　　Unpublished (2002)
COMMENT　　Other GSSs: OGWII71TH
　　　　　　Contact: Cathy Whitelaw
　　　　　　TIGR
　　　　　　9712 Medical Center Drive, Rockville, MD 20850, USA
　　　　　　Tel: 301-838-5843
　　　　　　Fax: 301-838-0208
　　　　　　Email: whitelaw@tigr.org
　　　　　　Seq primer: TF
　　　　　　Class: sheared ends.
FEATURES　　　　　　Location/Qualifiers
　　source　　　　　1..757
　　　　　　　　　　/organism="Zea mays"
　　　　　　　　　　/mol_type="genomic DNA"
　　　　　　　　　　/strain="B73"
　　　　　　　　　　/db_xref="taxon:4577"
　　　　　　　　　　/clone_lib="ZMMBMa0595K21"
　　　　　　　　　　/clone="ZM_0.7_1.5_KB"
　　　　　　　　　　/note="Vector: pBCSk-; Site_1: HincII; 0.7-1.5 kb
　　　　　　　　　　methylation filtered genomic DNA library"
ORIGIN

Query Match　　　3.0%;　　Score 73.6;　　DB 29;　　Length 757;
Best Local Similarity　46.6%;　　Pred. No.19;
Matches 306;　Conservative　0;　Mismatches　344;　Indels　6;　Gaps　2;

Qy　978 GAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGACGACGA 1037
　　　　　|| |||||||||||| |||| |||
Db　754 GCAGGTGGGCGTGAGCGACGTGTTCGCCGTGCCGGGGACTTCAACCTGACGCGCTGGA 695

Qy 1038 GGACTTCCGCAAGTACACCGCGCTTCACCATCCCAGCATCAACAACAG--ACCCCCGG 1094
　　　　　|| |||||| || ||| |||
Db 694 CCACCTGATCGCGGAGCCCGGCCTCCGCGTGTTCGGCTGCGCAAGGAGCTCAACGCCGG 635

Qy 1095 CATCCGCTACCAGTACACAACGTGCCCCCAGGGCTGGAAGGGAGCCCCAGCATCTTCCA 1154
　　　　　|||| |||| || |||
Db 634 GTACCGCCGCGACGGGTACGCGCGGGGCGTGGGGCGTGGGCGCCTGCGTTGTCACCTTCAC 575

Qy 1155 GAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCGGCCATCCAGCCGGAACCCCGAGATCGTGATCTA 1214
　　　　　||| ||||| |||
Db 574 CGTGGGCGGGCTCAGCGTGCTCACGCGTCCTCAACGCGGGCCAGGAGCGGAACCTGCCCGT 515

Qy 1215 CCAGTACATGGACGACCTTGACGTGGGCAGCGACCTGGAGATGGGCCAGGCACCGCGCCAA 1274
　　　　　||||||||| |||
Db 514 GGTGTGCGTGGGCGGGCGGGCCCAACTCCAAGCACTACGGCGTCCTCCACCA 455

Qy 1275 GATCGAGGAGCTGCGCAAGCACTGCACGTGCGCTGTCGGCTTCACCACCCCCGACAAGAAGCA 1334
　　　　　|| ||||| |||
Db 454 CACCATGGCGTGACCTCCGGACTTCTGCGAGGAGCTGCGCGGTGCGTTGGCTTCTGCCA 395

Qy 1335 CCAGAAGGAGCGCCCCCCTTCGCACGCGTCAGCAGTGGGCTACGAGTGCCACCAGTTGGACCGT 1394
　　　　　||| ||||| |||
Db 394 CCAGGCCGTCGTCAACAACCTGGAGGACGCGCACGA---GCAGATCGACGGGGCCATCTC 338

Qy 1395 GCAGCCATCGAGCTGCCCGAGAAGGAGAGCTGAACCGTGAACGACATCCAGAAGCTGGT 1454
　　　　　|||||| |||
Db 337 CACGGGACGCTCAGGGAGGAGCAAGCGGCCAGCCAGATGCTCACTGAGCTGCAACCTGCCGGGGGCT 278

Qy 1455 GGGCAAGCTGAACTGGGCCAGCAGCTGACGTGGGGATCTACCCCGGCATCAAGGTGCGCAGCTGTGCAA 1514
　　　　　|| ||||| |||
Db 277 GCCGCATCCGAGTTCGCCAGCCACCGGACCCTGCCATACTTGCTGCGCCAAGGCCTAGCAA 218

Qy 1515 GCTGCTGCGCCGCGCCAAGGCCGGCCAAGCCGCCGACATCGTGCCCCTGACCGGACCGGAGGAGGCCGAGCT 1574
　　　　　||| ||||| |||
Db 217 CCTGACGGGTCTCGCGCCGCCCGTCGGGCCGCCGCCACCGTGTCGTTCCTGACAAGGCCACCG 158

Qy 1575 GGAGCTGGCCGAGAACCGGAGATCGTGCGGCCGGTGCCCGTGCACGGCGTGTACTACG 1630
　　　　　|| ||||||||| |||
Db 157 GCCGGTGGTGTGGCCGGCCCCAGATCCGGTGCCCAAGATCCGGTGCCCAAGGCCAGGCCGGCCGCCGCCTCG 102

RESULT 15
BM321393
LOCUS　　　　　BM321393　　　　　853 bp　　mRNA　　　linear　　EST 03-JAN-2002
DEFINITION　　rockefeller 0.1222 Mastigamoeba balamuthi lambda ZAP II library
　　　　　　　　Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
　　　　　　　　sequence.
ACCESSION　　　BM321393
VERSION　　　　BM321393.1　　GI:18055799
KEYWORDS　　　　EST.
SOURCE　　　　　Mastigamoeba balamuthi
　ORGANISM　　　Mastigamoeba balamuthi
　　　　　　　　　Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE　　　1　(bases 1 to 853)
　AUTHORS　　　Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
　　　　　　　　Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
　　　　　　　　Philippe,H.
　TITLE　　　　The analysis of 100 genes supports the grouping of three highly
　　　　　　　　divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
　JOURNAL　　　Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
　MEDLINE　　　21819461
　PUBMED　　　　11830664
COMMENT　　　　Contact: Muller Miklos
　　　　　　　　Laboratory of Biochemical Parasitology
　　　　　　　　The Rockefeller University
　　　　　　　　1230 York Avenue, New York, NY 10021, USA
　　　　　　　　Email: mmuller@rockvax.rockefeller.edu
　　　　　　　　Insert Length: 853　Std Error: 0.00
　　　　　　　　POLYA=No.
FEATURES　　　　　　　　Location/Qualifiers
　　source　　　　　　　1..853
　　　　　　　　　　　　/organism="Mastigamoeba balamuthi"
　　　　　　　　　　　　/mol_type="mRNA"
　　　　　　　　　　　　/strain="ATCC 30984"
　　　　　　　　　　　　/db_xref="taxon:108607"
　　　　　　　　　　　　/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
　　　　　　　　　　　　/note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match　　　3.0%;　　Score 73.2;　　DB 12;　　Length 853;
Best Local Similarity　46.9%;　　Pred. No. 22;
Matches 299;　Conservative　0;　Mismatches　329;　Indels　9;　Gaps　2;

Qy　388 CTGTGGCAGCGCCCCTGGTGCAGATCAAGGTGGGCGGCAGATCAAGGAGGCCCTGCTG 447
　　　　　|| |||| |||
Db　125 CTGGCCCGCCGTGTGCTGAAGAAGAGCTCAACCTGGCAGGTGTCAAGAAG 184

Qy　448 GACACCGGCGCGACGACACCGTGCTGGAGGAGATGAGCCTGGCAAGTGCAAGTGGAAGCCC 507
　　　　　|| |||||| |||
Db　185 GTCAACGGCGGAGGACTACAACGTCGAGGAGCTGCACGCGCGCCGGGCCCGTTCAAGGCC 244

Qy　508 AAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATC 567
　　　　　|| |||||| |||
Db　245 CTGCTCGACGTCGGCGTCGGTCCGCACCTCGACCACCGGTGTCGTTCGGCCGCCCCTCAAG 304

Qy　568 GAGATCTGCGCAAGAAGGCCATCGGCCACCGGTGTGATGGCCCCACCCCGTGAACATC 627
　　　　　|| |||||| |||
Db　305 GGCATGTGCGACGGCGCGCGCGTCAACGTCCCC---CACAGCGAGACGCCTTCGTCGGCTTC 361

Qy　628 ATCGGCGGCAACGAAGACGAGCTCAACGCGCCGTCCTCCGCAAGTACATCTTCGGCGGCAC 687
　　　　　|| |||||| |||
Db　362 AACGGCGACAAGAAGAGACTCAACGCGGCCCGGACTGAGCTGAAGGTGAAGCAGTGGCC 421

Qy　688 GAGACCCGTGCCCGTGAAGTCATGGGGCATGACGCAGGACGCAGGAGGGCTAGCAA 747
　　　　　|| |||||| |||
Db　422 GTCGCCGCGTACATGAGCTCCTCAAGGAGCAGGAGCTCGTGCGACCGCCGCCAGTTC 481

```
Qy  748 CTGACCGAGGAGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAGGAGGAGGGC 807
            |||—|—|||—| ||| |||—||||— |||—|| ||—| |—| ||||—
Db  482 TCGCGCTACGCCAAGGAGGGTGTCACCGCCGACATGCTCGAGAAGATCTACACCGAGGCC 541

Qy  808 AAGATCACCAAGATCGGCCCCGAGAACCCCTACACACACCCCGTGTTCGCCATCAAGAAG 867
            ||||—|—||| |||—| |—| ||||—|||—|| |||—|||
Db  542 CA------CAAGCAGATCCGCGCCGACCGGACCTTCGTCCCCAAGCGGCCCTGAAGCCC 595

Qy  868 AAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAAACAAGCCACCCAG 927
            ||||| ||—||| ||—||||| |—| |||||||| —|
Db  596 GAGGGCGCCAAGCCCAAGCACTGGGCGCAAGCGCCAGGCTGACGTACCACCGAGCGCCAAGAAC 655

Qy  928 GACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCGCCGGCCTGAAGAAGAAGAAGAGAGC 987
            |—| |||—||| ||||| |||—|||||||| —|
Db  656 CGCGGTCGCCCAAGAAGGTCCGCCTGGGCTACCCCGACGCCCCCAAGAGACCCAGTAAATT 715

Qy  988 GTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCG 1024
            |||—| ||||—| |||—|||||| |||—|||
Db  716 CCCGGTGCCCTGTACACNGCCATCTCCCGCTGCGGCG 752
```

Search completed: April 10, 2004, 16:16:19
Job time : 4157.08 secs

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Run on: April 10, 2004, 02:53:16 ; Search time 6342.08 Seconds
(without alignments)
16873.640 Million cell updates/sec

Title: US-09-610-313-30
Perfect score: 2469
Sequence: 1 gtcgacgccaccatggccga........gggctagcaccggtgaattc 2469

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2469	100.0	2469	6	AX455914	AX455914 Sequence
2	2442.2	98.9	2463	6	AX455915	AX455915 Sequence
3	2415.4	97.8	2457	6	AX455916	AX455916 Sequence
4	2052	83.1	2312	6	BD263706	BD263706 Improved
5	2052	83.1	2312	6	AR373389	AR373389 Sequence
6	2051.2	83.1	9166	6	AX427930	AX427930 Sequence
7	2042.8	82.7	9788	6	AX427936	AX427936 Sequence
8	2028.8	82.2	9169	6	AX427931	AX427931 Sequence
9	2027	82.1	9194	6	AX427925	AX427925 Sequence
10	2027	82.1	12411	6	AX427927	AX427927 Sequence
11	2025.4	82.0	9194	6	AX427926	AX427926 Sequence
12	2025.2	82.0	2306	6	BD263704	BD263704 Improved
13	2025.2	82.0	2306	6	AR373387	AR373387 Sequence
14	2020.4	81.8	9785	6	AX427938	AX427938 Sequence
15	2018.6	81.8	9189	6	AX427921	AX427921 Sequence
16	2017	81.7	9167	6	AX427933	AX427933 Sequence
17	2017	81.7	9170	6	AX427928	AX427928 Sequence
18	2017	81.7	9782	6	AX427935	AX427935 Sequence
19	2017	81.7	9783	6	AX427934	AX427934 Sequence
20	2017	81.7	9792	6	AX427932	AX427932 Sequence
21	2015.6	81.6	4353	12	AF287353	AF287353 Synthetic
22	1999	81.0	3015	6	AX455946	AX455946 Sequence
23	1998.4	80.9	2300	6	BD263705	BD263705 Improved
24	1998.4	80.9	2300	6	AR373388	AR373388 Sequence
25	1982.4	80.3	3009	6	AX455987	AX455987 Sequence
26	1969.4	79.8	4319	6	BD263639	BD263639 Improved
27	1969.4	79.8	4319	6	AR373322	AR373322 Sequence
28	1958.8	79.3	4352	12	AF287352	AF287352 Synthetic
29	1907	77.2	2577	6	AX457088	AX457088 Sequence
30	1905.8	77.2	4343	6	AX306428	AX306428 Sequence
31	1901.8	77.0	4341	6	AX306429	AX306429 Sequence
32	1892.2	76.6	1977	6	AX455954	AX455954 Sequence
33	1879.2	76.1	7897	6	AX427923	AX427923 Sequence
34	1858	75.3	2305	6	BD263702	BD263702 Improved
35	1858	75.3	2305	6	AR373385	AR373385 Sequence
36	1837	74.4	1965	6	AX455952	AX455952 Sequence
37	1831.2	74.2	2299	6	BD263703	BD263703 Improved
38	1831.2	74.2	2299	6	AR373386	AR373386 Sequence
39	1825.6	73.9	3024	12	AF287355	AF287355 Synthetic
40	1807.6	73.2	9407	6	AX427937	AX427937 Sequence
41	1678.6	68.0	4307	6	AR302563	AR302563 Sequence
42	1678.6	68.0	4307	6	AX019132	AX019132 Sequence
43	1678.6	68.0	4307	6	BD130468	BD130468 Antiviral
44	1659.4	67.2	4307	6	BD268901	BD268901 Anti-vira
45	1659.4	67.2	4307	6	AX035453	AX035453 Sequence

ALIGNMENTS

RESULT 1
AX455914
LOCUS   AX455914   2469 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION Sequence 30 from Patent WO0204493.
ACCESSION   AX455914
VERSION   AX455914.1 GI:21714899
KEYWORDS   synthetic construct
SOURCE   synthetic construct
ORGANISM   artificial sequences.
REFERENCE   1
AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE   Polynucleotides encoding antigenic hiv type c polypeptides and uses thereof
JOURNAL   Patent: WO 0204493-A 30 17-JAN-2002;

```
FEATURES    CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
source      1..2469
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:3630"
            /note="PR975(+)"
ORIGIN

Query Match       100.0%;  Score 2469;  DB 6;  Length 2469;
Best Local Similarity  100.0%;  Pred. No. 4.6e-250;
Matches 2469;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

```
Qy     1 GTCGACGCCACCATGGCCGAGGCCATGGCCGCCAGGCCACCAGCGGCCAACATCCTGATGCAG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GTCGACGCCACCATGGCCGAGGCCATGGCCGCCAGGCCACCAGCGGCCAACATCCTGATGCAG 60

Qy    61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120

Qy   121 CACATCGCCCCAACTGCGCGGGCCCCCGCAAGAAGAGGCCAGGCCAACTGCGCAAGGAGG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACATCGCCCCAACTGCGCGGGCCCCCGCAAGAAGAGGCCAGGCCAACTGCGCAAGGAGG 180

Qy   181 GGCCACCAGATGAAGGACTGCACCGAGGACTGCTGGGCAAGCTTCCGGCGGGACCTGGCC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCCACCAGATGAAGGACTGCACCGAGGACTGCTGGGCAAGCTTCCGGCGGGACCTGGCC 240

Qy   241 TTCCCCCAGGGCAAGCCGGAGTTCCCCGAGTTCCGGCGGGACAACCGCCAACAGCCCACC 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCCCCAGGGCAAGCCGGAGTTCCCCGAGTTCCGGCGGGACAACCGCCAACAGCCCCACC 300

Qy   301 AGCCGGAGCTGCAGGTGCGGGCGACAACCCCCAGCGAGGCGGCGGCCGAGCGCGCCAG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCCGGAGCTGCAGGTGCGGGCGACAACCCCCAGCGAGGCGGCGGCCGAGCGCGCCAG 360

Qy   361 GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTG 420

Qy   421 GGCGGGCAGATCAAGAGAGGCCCTGCTGGACACCGGCGCGGACGACACCGGTGCGCACGG 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGGCAGATCAAGAGAGGCCCTGCTGGACACCGGCGCGGACGACACCGGTGCGCACGG 480

Qy   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAG 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAG 540

Qy   541 GTGTGGCCAGTACGACCCGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGTGGCCAGTACGACCCGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG 600

Qy   601 CTGATCGGCGCCCACCCCCGTGAACATATCGGCCGCAACATGTCGACCCAGCTGGGCTGC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCGCCCACCCCCGTGAACATCATCGGCCGCAACATGTCGACCCAGCTGGGCTGC 660

Qy   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720

Qy   721 GACGGGCCCCAAGGTGAAGCAGTGGGCAAGAGGAGAGATCAAGGCCCTGACCGCC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGGCCCCAAGGTGAAGCAGTGGGCAAGAGGAGAGATCAAGGCCCTGACCGCC 780

Qy   781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840

Qy   841 AACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
```

```
Qy   901 TTCCCGGAGCTGACAACAGCCACCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCAC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCCGGAGCTGACAACAGCCACCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCAC 960

Qy   961 CCCGCCGGCCTGAAGAAGAGAGAGAGCTTCGGACAGTACACCACCGCCTTCACCATCAAC 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCCGGCCTGAAGAAGAGAGAGAGCTTCGGACAGTACACCACCATCCCCAGCATCAAC 1020

Qy  1021 AGCGTGCCCCTGGACGAGGACTTCCGACAGTACGAAGTACACCGGCCTTCACCATCAAC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AGCGTGCCCCTGGACGAGGACTTCCGACAGTACGAAGTACACCGGCCTTCACCATCAAC 1080

Qy  1081 AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140

Qy  1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC 1200

Qy  1201 GAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGC 1260

Qy  1261 CAGCACCGGCCAAGATCGAGGAGTCGGCAAGCACCTGCTGCGGGCTTCACCACC 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CAGCACCGGCCAAGATCGAGGAGTCGGCAAGCACCTGCTGCGGGCTTCACCACC 1320

Qy  1321 CCCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGTCGGATGGCTACGAGCTGCACCCC 1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CCCGACAAGAAGCACCAGAAGGAGCACCCCTTCCTGTGTGATGGCTACGAGCTGCACCCC 1380

Qy  1381 GACAAGTGACCGTGCAGCCATCGAGCTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGGTG 1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 GACAAGTGACCGTGGTGGCAGCTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGGTG 1440

Qy  1441 ATCCAGAGCTGGTGCGGGAAGCTGAACTGGGCCAAGGCCCTGACCGACATCGTGCCCTGACC 1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 ATCCAGAGCTGGTGCGGGAAGCTGAACTGGGCCAAGGCCCATCACCCGGCATCGTGCCCTGACC 1500

Qy  1501 CGCCAGCTGTTGCAAGCTGCTGCGGCGGCGGCCAGGAACCGGCAGAATCCTGCGGGAGCCCGTGCACGGC 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CGCCAGCTGTGCAAGCTGCTGCGGCGGCGGCCAGGAACCGGCAGAATCCTGCGGGAGCCCGTGCACGGC 1560

Qy  1561 GAGGAGCCCAGCTGGACGTGGCCGGCCTGGAGCTGGAGCGGCCATCGAGGCGCATCAAG 1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GAGGAGCCCAGCTGGACGTGGCCGGCCTGGAGCTGGAGCGGCCATCGAGGCGCATCAAG 1620

Qy  1621 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCAAGAAGCAGGGCCACGACCAG 1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCAAGAAGCAGGGCCACGACCAG 1680

Qy  1681 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACTGAAGACCGGCAAGTACGCCCAAG 1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACTGAAGACCGGCAAGTACGCCCAAG 1740

Qy  1741 ATGGCAGCGCCCACACCAACGCGTGACCGAGGCCGTGCAGAAGATCGCC 1800
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ATGGCAGCGCCCACACCAACGCGTGACCGAGGCCGTGCAGAAGATCGCC 1800

Qy  1801 ATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 ATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1860

Qy  1861 ACCTGGGAGACCTGGTGGACCTACTGGCAGGCCACCTGCGATCCCGAGTGGGGAGTTC 1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACCTGGGAGACCTGGTGGACCTACTGGCAGGCCACCTGCGATCCCGAGTGGGGAGTTC 1920

Qy  1921 GTGAACACCCCCGTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC 1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 GTGAACACCCCCGTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC 1980

Qy  1981 GCCGAGACCTTCAGTGACGTGGACGCGCCCAACCCGCAGAGACCGAAGATCGGCAAGGCCCGGC 2040
```

```
Db   1981  GCCGAGACCTTCTACGTGACGCGGGGCGCCGGCGGGGGCGCCGGCGGGGGCGCCGGGGCAAGATCGGCCAAGATCGGCCAAGGCCGGC  2040
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2041  TACGTGACCGACCGGCGGGGCGGGGCGGCGGCGAGAAGATCGTGAGCCTGACGGAGCTGACCGAGACCAACCAGAAG  2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db   2041  TACGTGACCGACCGGCGGGGCGGCGGCGAGAAGATCGTGAGCCTGACCGAGACCTGAGACCAACCAGAAG  2100

Qy   2101  ACCGAGCTGCAGGCCATCCAGTGGCCCTGAGCAGGCGGCAGCGGCAGCAGGTGAACATCGTG  2160
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101  ACCGAGCTGCAGGCCATCCAGTGGCCCTGGGCCCGCGACGAGCAGCGGCAGCGGCAGCATCGTG  2160

Qy   2161  ACCGACAGCCAGTACGGCCTGGCCATCATCCAGGCCCAGCGACCGGCGAGAGCGAGAGCGGAG  2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161  ACCGACAGCCAGTACGGCCTGGCCATCATCCAGGCCCAGCAAGAGCGCAAGAGCGAGAGCCGAG  2220

Qy   2221  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG  2280
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG  2280

Qy   2281  CCCGCCCACAAGGGCATCGGCGGCAGCGGCATCGACACGAGCTGGTCAAGGGCATC  2340
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281  CCCGCCCACAAGGGCATCGGCGGCAGCGGCATCGACACAGAGCTGGTGAGCAAGGGCATC  2340

Qy   2341  CGCAAGGTGCGTGTTCCTGGACGGCGGCGGGCCCTAGGATGGCGGCCATCGTGATCTACCAGTACAATGGAC  2400
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341  CGCAAGGTGCGTGTTCCTGGACGGCGGCGGGCCCTAGGATGGCGGCATCGTGATCTACCAGTACAATGGAC  2400

Qy   2401  GACCTGTACGTGGGCAGCAGCGGCGGCGGCGGCCCTAGGATCGGATGGATTAAAAGCTTCCCGGGGCTAGCACC  2460
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401  GACCTGTACGTGGGCAGCAGCGGCGGCGGCGGCGGCCCTAGGATCGGATCGGATTAAAAGCTTCCCGGGGCTAGCACC  2460

Qy   2461  GGTGAATTC  2469
             |||||||||
Db   2461  GGTGAATTC  2469


RESULT 2
AX455915
LOCUS       AX455915              2463 bp    DNA         linear      PAT 06-JUL-2002
DEFINITION  Sequence 31 from Patent WO0204493.
ACCESSION   AX455915
VERSION     AX455915.1   GI:21714900
KEYWORDS
SOURCE      synthetic construct
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1
AUTHORS     zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE       Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
JOURNAL     Patent: WO 0204493-A 31 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES             Location/Qualifiers
source          1..2463
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="PR975YM"
ORIGIN

Query Match    98.9%;  Score 2442.2;  DB 6;  Length 2463;
Best Local Similarity  99.6%;  Pred. No. 3e-247;
Matches 2460;  Conservative  0;  Mismatches  3;  Indels  6;  Gaps  1;

Qy      1  GTCGAGACGCCACCATGGCCGAGGCCATGAGGCCAGGCCACCAGCCCAACATCCTGATGCAG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GTCGAGACGACCCCGGCCATGGCCGAGGCCATGAGGCCAGGCCACCAGCCCAACATCCTGATGCAG  60

Qy     61  CCCAGCAACTTCAAGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CGCAGCAACTTCAAGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120

Qy    121  CACATCGCCCGCAACTGCCCCGCGCCCCCCCGCAAGAAGAGGGCTGCTGGAAGTGCGGCAAGGAG  180
```

```
Db    121  CACATCGCCCGCAACTGCCCGCGGCCCCCCCGCGAGATCGGCCAAGATCGGCCAAGGAGGAG  180

Qy    181  GGCCACCAGATGAAGGACTGACCGAGAAGATCGTGAGCCTGACGGAGCTGACCTGGCC  240
             |||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db    181  GGCCACCAGATGAAGGACTGACCGAGAAGATCGTGAGCCTGACCTGAGCTGACCTGGCC  240

Qy    241  TTCCCCCAGGGCAAGGCCCCGAGTTCCCCAGCGAGGAGCAGAAACCGGCCAACAGCCCCACC  300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  TTCCCCCAGGGCAAGGCCCCGAGTTCCCCAGCGAGGAGCAGAGAACCGGCCAACAGCCCCACC  300

Qy    301  AGCCGGCAGTGCAGGTGCGGCGGCAACCCCAGCCGGCCGGCCGCGGCCGGCCGAGCGCCAG  360
             |||||||||||||||||||||||||||||||||||||||||||||  ||||
Db    301  AGCCGGCAGTGCAGGTGCGGCGGCGACCAACCCCAGCGAGGCCGGCCGCGGCCGGCCGAGCGCCAG  360

Qy    361  GGCACCCTGAACTTCCCCAGGATCACCCGTGTGGAGCGCGCCCCTGTGAGCATCAAGGTG  420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  GGCACCCTGAACTTCCCCAGCATCACCCGTGTGGCAGCGCGCCCCCTGGTGAGCATCAAGGTG  420

Qy    421  GGCGGCAGATCAAGAGCCCTGCTGGACAACGGCAGCGAGATCGACAAGGGCCATC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GGCGGCAGATCAAGAGCCCTGCTGGACAACGGCAGCGAGATCGACACAGAGCGGCATC  480

Qy    481  ATGAGCTGCCCGGCAGTGGTGAAGCCCAAGAGTGAAGCCCGATCGTGCGGCAAGAGATGATCGGCCACACCGTG  540
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  ATGAGCTGCCCGGCAGTGGTGAAGCCCAAGAGTGAAGCCCGATCGTGCGGCAAGAGATGATCGGCCACACCGTG  540

Qy    541  GTGCGCAGTACGACCCACCCCCGTGAACATCATCGGCCGCCAACATGTCGACCAGCTGGCTGC  600
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GTGCGCAGTACGACCCACCCCCGTGAACATCATCGGCCGCCAACATGTCGACCAGCTGGCTGC  600

Qy    601  CTGATCGGCCCGGGCCCGTGCCCCATCAGCCCATCCCCGTGAAGCTGAAGCTGAAGCGCCGGCATG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CTGATCGGCCCGGGCCCGTGCCCCATCAGCCCATCCCCGTGAAGCTGAAGCTGAAGCGCCGGCATG  720

Qy    721  GACGGCGCCCAAGGTGAAGCAGTGGCCCCTGACGGAGGAGAGGAGATCAAGGGCCCTGACCGGCGC  780
             ||||||||||||||||||||||||||||||||||||||||||||  ||||
Db    721  GACGGCGCCCAAGGTGAAGCAGTGGCCCCTGACCGGGAGAAGATCAAGGGCCCTGACCGGCGC  780

Qy    781  ATCTGGCAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC  840
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  ATCTGGCAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC  840

Qy    841  AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCAGCAAGTGGCGCAAGCTGGTGGAC  900
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCAGCAAGTGGCGCAAGCTGGTGGAC  900

Qy    901  TTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC  960
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC  960

Qy    961  CCCGCCGGCCTGAAGAAGAAGAAGGCCGTGACGGCGCTGACCGACGTGGGGGCACGCTACTTC  1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  CCCGCCGGCCTGAAGAAGAAGAAGGCCGTGACGGCGCTGACCGACGTGGGGGCACGCTACTTC  1020

Qy   1021  AGCGTGCCCCTGGACGGAGGACTTCCGCAAGTAGTACCGGCTTCACCATCCCCAGCATCAAC  1080
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  AGCGTGCCCCTGGACGGAGGACTTCCGCAAGTAGTACCACCGGCTTCACCATCCCCAGCATCAAC  1080

Qy   1081  AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140

Qy   1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGAGATCCGCGCCCGCAAGCCCC  1200
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTTCAACTGCGGCAAGCCCC  1200

Qy   1201  GAGATCGTGATCTACCAGTACATGGACGACATCTGTACGTGGGCAGCGACCTGGAGATCGGC  1260
             |||||||||||||||||||        ||||  |||||||||||||||||
```

```
Qy   2341 CGCAAGGTGCTGTTCCTGGACGGCGCATCGATGGGCATCGTGATCTACCAGTACACATGGAC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2335 CGCAAGGTGCTGTTCCTGGACGGCGCATCGATGGGCATCGTGATCGTGATCTACCAGTACACATGGAC 2394

Qy   2401 GACCTGTACGTGGGCAGCGGCCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2395 GACCTGTACGTGGGCAGCGGCCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC 2454

Qy   2461 GGTGAATTC 2469
          |||||||||
Db   2455 GGTGAATTC 2463
```

RESULT 3
LOCUS   AX455916   2457 bp   DNA   linear   PAT 06-JUL-2002
DEFINITION Sequence 32 from Patent WO0204493.
ACCESSION   AX455916
VERSION   AX455916.1 GI:21714901
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
  TITLE   Polynucleotides encoding antigenic hiv type c polypeptides,
            polypeptides and uses thereof
  JOURNAL   Patent: WO 0204493-A 32 17-JAN-2002;
            CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES   Location/Qualifiers
     source   1..2457
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PR975MWM"

ORIGIN

Query Match   97.8%; Score 2415.4; DB 6; Length 2457;
Best Local Similarity   99.3%; Pred. No. 1.9e-244;
Matches 2451; Conservative   0; Mismatches   6; Indels   12; Gaps   2;

```
Qy      1 GTCGACGGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCGACGGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60

Qy     61 CGCAGCAACTTCAAGGGCCCCATCAGTGCTTCAACTGCGGCAAGGAGGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCAGCAACTTCAAGGGCCCCATCAGTGCTTCAACTGCGGCAAGGAGGGC 120

Qy    121 CACATCGCCCCAACTGCCGCGCCCCCGCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACATCGCCCCAACTGCCGCGCCCCCGCAAGAAGGGCGTGCTGGAAGTGCGGCAAGGAG 180

Qy    181 GGCCACCAGATGAAGGACTGAAGGACTGACCGCAGAGCGAGCAACTTCTTCCGGCAACCAGCCC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCCACCAGATGAAGGACTGACCGCAGAGCGAGCAACTTCTTCCGGCAACCAGCCC 240

Qy    241 TTCCCCCAGGGCGAAGGCCCGGCAGTTCCCGGCGACAACCCGCAGAGCCCCACC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCCCCCAGGGCGAAGGCCCGGCAGTTCCCGGCGACAACCCGCAGAGCCCCACC 300

Qy    301 AGCCGCGAGCTGCAGGTGCGCGGCGGCGCAGCAACCCCGAGCGGCCAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCCGCGAGCTGCAGGTGCGCGGCGGCGCAGCAACCCCGAGCGGCCAG 360

Qy    361 GGACACCCTGAACTTCCCCCAGATCACCGTGGGCGACATCAAGGTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGACACCCTGAACTTCCCCCAGATCACCGTGGGCGACATCAAGGTG 420

Qy    421 GGGGGCCAGATCAAGGAGGCCCTGCTGGACACCGGGCGCGACACCGGCGCGACACCGTGTGGGAGGAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGGGGCCAGATCAAGGAGGCCCTGCTGGACACCGGGCGCGACACCGGCGCGACACCGTGTGGGAGGAG 480
```

```
Db   1201 GAGATCTGATCTACCA------GGCCCCTCGTACGTGGGCAGCGACCTGGAGATCGGC 1254
Qy   1261 CAGCACCCGCGCCAAGATCGAGGAGCTGCGCGCAGCACCTGCTGCGCGTGGGGCTTCACCACC 1320
Db   1255 CAGCACCCGCGCCAAGATCGAGGAGCTGCGCGCAGCACCTGCTGCGCGTGGGGCTTCACCACC 1314
Qy   1321 CCGACAAGAAGCACCAGAAGCACCAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCC 1380
Db   1315 CCGACACAAGAAGCACCAGAAGCACCAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCC 1374
Qy   1381 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC 1440
Db   1375 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC 1434
Qy   1441 ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCGGCCATCAAGGTG 1500
Db   1435 ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCGGGCATCAAGGTG 1494
Qy   1501 CGCCAGCTGTGCAAGCTGCTGCGCGGCGGCAAGGCCCTGACCGACATCGTGCCCCTGACC 1560
Db   1495 CGCCAGCTGTGCAAGCTGCTGCGCGGCGGCAAGGCCCTGACCGACATCGTGCCCCTGACC 1554
Qy   1561 GAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGC 1620
Db   1555 GAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGC 1614
Qy   1621 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAG 1680
Db   1615 GTGTACTACGACCCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAG 1674
Qy   1681 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGCAAG 1740
Db   1675 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGCAAG 1734
Qy   1741 ATGCGCACCGCCCACACCAACCACCCCAACGAGCGTGAAGCAGCAGTGACCGAGAAGATCGCC 1800
Db   1735 ATGCGCACCGCCCACACCAACCACCCCAACGAGCGTGAAGCAGCAGTGACCGAGAAGATCGCC 1794
Qy   1801 ATGGAGAGCACCTGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAGGGAG 1860
Db   1795 ATGGAGAGCACCTGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1854
Qy   1861 ACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC 1920
Db   1855 ACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC 1914
Qy   1921 GTGAACACCCCCCTGTGTGAAGCTGTGGTACAGCTGGAGAAGAGAGCCCATCATCGGC 1980
Db   1915 GTGAACACCCCCCTGTGTGAAGCTGTGGTACAGCTGGAGAAGAGAGCCCATCATCGGC 1974
Qy   1981 GCCGAGACCTTCTACGTGGACGGCGGCCGCCAACCGGCCAAGATCGGCAAGGCCGGC 2040
Db   1975 GCCGAGACCTTCTACGTGGACGGCGGCCGCCAACCGGCCAAGATCGGCAAGGCCGGC 2034
Qy   2041 TACGTGACCGACCGGGCCGCCGGCCGCGAAGATCGTGACCGAGACCGACACCAACCAGAAG 2100
Db   2035 TACGTGACCGACCGGGCCGCCGGCCGCGAAGATCGTGACCGAGACCGACACCAACCAGAAG 2094
Qy   2101 ACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGAGTGAACATCGTG 2160
Db   2095 ACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGAGTGAACATCGTG 2154
Qy   2161 ACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGGAG 2220
Db   2155 ACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGGAG 2214
Qy   2221 CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG 2280
Db   2215 CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG 2274
Qy   2281 CCCGCCCACAAGGGCATCGGCGGCAAGATCGAGCGGCAAGATCGAGCGGCCAACGAGCAAGGGCATC 2340
Db   2275 CCCGCCCACAAGGGCATCGGCGGCAAGATCGAGCGGCAACGAGCAAGGGCATC 2334
```

```
Qy   1561  GAGGAGGCCGAGCTGGAGCTGCCGAGAACCGCCGAGATCCTGCGCGAGCCGTGCACGGC  1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1549  GAGGAGGCCGAGCTGGAGCTGCCGAGAACCGCCGAGATCCTGCGCGAGCCGTGCACGGC  1608

Qy   1621  GTGTACTACGACCCCAGCAAGGAGACCTGGTGCGCAGACCAGAGCAGGCCCACGACCAG  1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1609  GTGTACTACGACCCCAGCAAGGAGACCTGGTGCGCAGATCAGAGCAGCAGGGCCACCACCAG  1668

Qy   1681  TGGACCTACCAGATCTACCAGGAGAGCCCTTCAAGACCTGAAGACCTGAGAGTACGGCCAAG  1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1669  TGGACCTACCAGATCTACCAGGAGAGCCCTTCAAGACCTGAAGACCTGAAGACCAGTACGGCCAAG  1728

Qy   1741  ATGGCGCACGGCCACCACCGACGTGAAGCAGTGAAGCTGACCGAGGCCGTGAGAGAATCGCC  1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1729  ATGGCGCACCGCCACCACCGACGTGAAGCAGTGAAGCTGACCGAGGCCGTGCAGAAGATCGCC  1788

Qy   1801  ATGGAGCAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCCAGTTCCGCCCATCCAGAAGGAG  1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1789  ATGGAGACGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCCAGTTCCGCCCATCCAGAAGGAG  1848

Qy   1861  ACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC  1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1849  ACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC  1908

Qy   1921  GTGAACACCCCCCCTGGTGAAGTGTGTACCAGCTGAGAAGGAGAGAGCCCAAGATCGGCAAGGCCGGC  1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1909  GTGAACACCCCCCCTGGTGAAGTGTGTACCAGCTGAGAAGGAGAGAGCCCAAGATCGGCAAGGCCGGC  1968

Qy   1981  GCCGAGACCTTCTACGTGACCGACCGGCGGCGAGAAGATCGTGAGCCTGACCGAGACCAACCAGAAG  2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1969  GCCGAGACCTTCTACGTGACCGACCGGCGGCGAGAAGATCGTGAGCCTGACCGAGACCAACCAGAAG  2028

Qy   2041  TACGTGACCAGATCATCGAGCAGCTGATCAAGGAGAAGGAGGTGTACCTGAGCTGGGTG  2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2029  TACGTGACCAGATCATCGAGCAGCTGATCAAGGAGAAGGAGGTGTACCTGAGCTGGGTG  2088

Qy   2101  ACCGAGTGCAGGCCATCCAGCTGCCCTGGCCCAGCGGCAGCGAGGTGAACATCGTG  2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2089  ACCGAGTGCAGGCCATCCAGCTGCCCTGGCCCAGCGGCAGCGAGGTGAACATCGTG  2148

Qy   2161  ACCGAGCAGCAGTACGCCCTGGGCATCATCCAGGCCCAGAACAGAGGGCGAGAGCGAG  2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2149  ACCGAGCAGCAGTACGCCCTGGGCATCATCCAGGCCCAGAACAGAGGGCGAGAGCGAG  2208

Qy   2221  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGGAGAAGGAGGTGTACCTGAGCTGGGTG  2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2209  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGGAGAAGGAGGTGTACCTGAGCTGGGTG  2268

Qy   2281  CCCGCCCACAAGGGCATCGGCAACGAGCATCGGCGAACAAGCTGGTGACAAGGGCATC  2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2269  CCCGCCCACAAGGGCATCGGCAACGAGCATCGGCGAACAAGCTGGTGACAAGGGCATC  2328

Qy   2341  CGCAAGGTGCTGTTCCTGGACGGCATCGATGGGCGCATCGATGTCGTGATCTACCAGTACATGGAC  2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2329  CGCAAGGTGCTGTTCCTGGACGGCATCGATGCGGCATCGATGTCGTGATCTACCAGTACATGGAC  2388

Qy   2401  GACCTGTACGTGGGCAGCCGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC  2460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2389  GACCTGTACGTGGGCAGCCGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC  2448

Qy   2461  GGTGAATTC  2469
           |||||||||
Db   2449  GGTGAATTC  2457
```

```
Qy    481  ATGAGCCTGCCCGCCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  ATGAGCCTGCCCGCCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG  540

Qy    541  GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCCAAGAGGCCATCGGCCACCGTG  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCCAAGAGGCCATCGGCCACCGTG  600

Qy    601  CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  CTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC  660

Qy    661  ACCCTGAACTTCCCATCAGCCCATCGAGACCCATGCCGTGAAGCTGACCCGCCGGCATG  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  ACCCTGAACTTCCCATCAGCCCATCGAGACCCATGCCGTGAAGCTGACCGCCGGCATG  720

Qy    721  GACGGCCCCAAGTGAAGCAGTGCCCCTGACCGAGAAGAATCAAGGCCCCTGACCGCC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GACGGCCCCAAGTGAAGCAGTGCCCCTGACCGAGAAGAATCAAGGCCCCTGACCGCC  780

Qy    781  ATCTGCGAGGAGATGGAGAAGGAGGGCCAAGATCACCAAGATCGCCCCGAGACCCCTAC  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  ATCTGCGAGGAGATGGAGAAGGAGGGCCAAGATCACCAAGATCGCCCCGAGAACCCCTAC  840

Qy    841  AACACCCCGTGTTCGCCATCAAGAAGGACGACAGCACCAAGTGGCGCAAGTGGTGGTGAC  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  AACACCCCGTGTTCGCCATCAAGAAGGACGACAGCACCAAGTGGCGCAAGTGGTGGTGAC  900

Qy    901  TTCCGGGAGCTGCAGAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTCCGGGAGCTGCAGAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC  960

Qy    961  CCCGCCGGCCTGAAGAGAAGAGAAGAGAAGAGGACTTCCGACGTACACCGCCTTCACCAGATCAAC  1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  CCCGCCGGCCTGAAGAGAAGAGAAGAGAAGAGGACTTCCGACGTACACCGCCTTCACCAGATCAAC  1020

Qy   1021  AGCGTGCCCCTGACGGAGACTTCGCGCAAGTACACCGCCTTCACCAGATCCCAGCATCAAC  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  AGCGTGCCCCTGACGGAGACTTCGCGCAAGTACACCGCCTTCACCAGATCCCAGCATCAAC  1080

Qy   1081  AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCCAGC  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140

Qy   1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCC  1200

Qy   1201  GAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGC  1260
           ||||||||||||||          |||||||||||||||||||||||||||||||||||||
Db   1201  GAGATCGTGATCTACCA------GGCCCCCGTGTACGTGGGCAGCGACCTGGAGATCGGC  1254

Qy   1261  CAGCAGAAGATCGAGGAGTGCGCCAAGACCTGCTGCGCGGGCTTCACCACC  1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255  CAGCAGAAGATCGAGGAGTGCGCCAAGACCTGCTGCGCGGGCTTCACCACC  1314

Qy   1321  CCCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGTGATGGCTACGACCCTGCACCCC  1380
           |||||||||||||||||||||||          |||||||||||||||||||||||||
Db   1315  CCCGACAAGAAGCACCAGAAGGAGCCC-----CGAGCTGCAGCCGCCCTGACC  1368

Qy   1381  GACAAGTGGACCGTGCAGCCCATCGAGAGCTGCCCGAGGAAGAGAGCTGGACCGTGAACGAC  1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1369  GACAAGTGGACCGTGCAGCCCATCGAGAGCTGCCCGAGGAAGAGAGCTGGACCGTGAACGAC  1428

Qy   1441  ATCCAGAGCTGGTGGGCAACGTGGGCAACTGAACTGGCCAGCTGAACTCACCCGGCATCAAGGTG  1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1429  ATCCAGAGCTGGTGGGCAACGTGGGCAACTGAACTGGGCCAGCCAGAATCTACCCCGGCATCAAGGTG  1488

Qy   1501  CGCCAGCTGTGCAAGGTGCTGCTGCGCGCGGCGCCAAGGGCCCTGACCGTGCCCCTGACC  1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1489  CGCCAGCTGTGCAAGGTGCTGCTGCGCGCGGCGCCAAGGGCCCTGACATCGTGCCCCTGACC  1548
```

RESULT 4
BD263706
LOCUS   BD263706   2312 bp   DNA   linear   PAT 17-JUL-2003
DEFINITION   Improved expression of HIV polypeptides and production of
             virus-like particles.
ACCESSION   BD263706
VERSION   BD263706.1   GI:33073474
KEYWORDS   JP 2002533124-A/73.

SOURCE   synthetic construct
ORGANISM   synthetic construct
            artificial sequences.
REFERENCE 1 (bases 1 to 2312)
AUTHORS   Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
            Greer,C., Selby,M. and Walker,C.
TITLE   Improved expression of HIV polypeptides and production of
            virus-like particles
JOURNAL   Patent: JP 2002533124-A 73 08-OCT-2002;
            CHIRON CORP
COMMENT   OS   Artificial Sequence
            PN   JP 2002533124-A/73
            PD   08-OCT-2002
            PF   30-DEC-1999 JP 2000591193
            PR   31-DEC-1998 US 60/114495, 01-DEC-1999 US 60/168471 PI
            SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
            KARIN HARTOG,
            PI   HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
            C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
            C12N5/10,
            PC   C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02,A61K37/02 CC
            Description of Artificial Sequence: FS(-).protmod.RTopt(+) FH Key
FEATURES   FT   source   1..2312
            FT   /organism='Artificial Sequence'.
    source   Location/Qualifiers
            1..2312
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"

ORIGIN

Query Match   83.1%; Score 2052; DB 6; Length 2312;
Best Local Similarity 93.6%; Pred. No. 2.3e-206;
Matches 2165; Conservative 0; Mismatches 135; Indels 12; Gaps 2;

Qy   170 GCGGCAAGGAGGGCCACCAGATGAAGGACTGACCGAGCGCCAGGCCAACTTCTTCCGCG 229
        |||||||||||||||||||| ||||| |||||| |||||| ||||||||||||||||||
Db   1 GCGGCCGCGGAAGGACACCAAATGAAAGATTGCACTGAGAGACCAGCAGCTTAATTTCTTCCGCG 60

Qy   230 AGGAACCTGCCTTCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGGCCAGAACCGGCGCCA 289
        |||||||| ||||||||||||||||||||||||| |||||||||||||||||||||||
Db   61 AGGAACCTGCCTTCCTGCAGGGCAAGGCCCGCGAGTCAGCAGGCAGAGACCGGCGCCCA 120

Qy   290 ACAGCCCCACCAGCGCGGCGAGCTGCAGTGCCGCG------ACAACCCCGCCAGCGAGG 343
        || ||||||||||||||||||||||||||||||||| ||||
Db   121 ACAGCCCCACCAGCGCGGCGAGCTGCAGGTGTGGGGCGGGAGAACAACAGCCTGGAGCCAGG 180

Qy   344 CCGGCGCCGAGCGCCAGGGCACCTG-----AACTTCCCCAGATCACCCTGTGGCAGC 397
        ||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   181 CCGGCGCCGACCGCCAGGGCACCGGAGCTTCAACTTCCCCAGATCACCCTGTGGCAGC 240

Qy   398 GCCCCCTGGTGAGACATCAAGGTGCGGGCAGATCGGCGGCCTGGCGGCGCACCGGCG 457
        ||| |||||||||||||||||| |||||||||||||| |||||||||||||||
Db   241 GCCCCCTGGTGACCATCGGCGAGGATGAGGAGATCGGCGGCCTGGCGGCTGCTGCACACCGGCG 300

Qy   458 CCGACGACACCGGTGCTGAGGAGATGAGCCTGCCCGGCAAGTGAAGCCAAGATGATCG 517
        ||||||||||| |||||||||||||||||||||||||| ||||||||||||||
Db   301 CCGACGACACCGGTGCTGAGGAGATGAGCCTGCCCGGCAAGTGAAGCCAAGATGATCG 360

Qy   518 GCGGCATCGGCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCG 577
        |||||||||||||||||||||||||||||||| |||||||||||| ||||||||
Db   361 GCGGGATCGGCGGCGGCTTCATCAAGGTGCGCAGTACGACCAGATCCCCGTGGAGATCTGCG 420

Qy   578 GCAAGAAGGCCATCGGCACCGTGCTGATCGGCGGCCGTGCCCCACCCGTGAACATCATCGGCCGCA 637
        |||||||||||||||| |||||||| |||||||||||||| |||||||||||||
Db   421 GCCACAAGGCCATCGGCACCGTGCTGGTGGTCGGCGGCCGTGCCCCCGTGAACATCATCGGCCGCA 480

Qy   638 ACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCC 697
        |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   481 ACCTGCTGACCCAGCTGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCC 540

Qy   698 CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG 757

```
Db   1621  CCGAGGCGCTGCGAGAAGTGAGCACCCAGAGAGCATCGTGATCTGGGGACCGACTACTGGCAGCCAAGT 1680
                 |||||   ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Qy      1  GCGGGCCGGGAAGCGACCACCCAATGAAAGATTGCACTGAGGACAGCCTAATTCTTCCGCG    60

Qy   1838  TCCGCCTGCCCATCCAGAAGAGACCTTGGGAGACCTTGGTGGACCACCTGGCAGGCCA 1897
               | ||||||||||||||||||||||||||||||||||||| ||||||| |||||||
Db    230  AGGACCTGGCCTTCCCCAGGCGCCTTCCCGAGGTTCCCAGCGAGACGAGAACCGCGCCA   289

Db   1681  TCAAGCTGCCCATCCAGAAGAGACCTTGGGAGTTGTGAACACACCCCCCCCCCCCGTTCAGCAGCGACGAGCAGCGAGCGG 1740
               |||||||||||||||||||||||||||||  |||||||||||||||||||||  |||||||
Qy     61  AGGACCTGGCCTTCCTGCGAGGCCCAAGGCCCAAGGCCCAGTTCAGCAGCGAGCGAGCGAGACCCGCGCCA   120

Qy   1898  CCTTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGGTGAAGCTGTGGTACCAGC 1957
               |||||||||||||||||||||||||||||||||      ||||| ||||| |||||||
Db    290  ACAGCCCACCAGCCGCGAGCTGCGAGGTGCGGGCGG------ACAACCCCGCAGCGAGG   343

Db   1741  CCTTGGATCCCCGAGTGGGAGTTCGTGAACACACCCCCCTGTGAAGCTGTGGTACCAGC 1800
               |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Qy    121  ACAGCCCACCAGCCGCGAGCTGCGAGGTGCGGGCGGCGGCGGCGGCAACAGCCTGAGCGAGG   180

Qy   1958  TGGAGAGGAGGCCCATCATCGGGCGGCCGGAGACCTTCTACGTGGACGGCGGCCAACCGG 2017
               |||||||||||||||||||  ||||| ||||| ||  AACTTCCCCAGATCACCCTGTGCAGC   397
Db    344  CCGGGCCGGGAGCGCCAGGGCACCCTG------AACTTCCCCAGATCACCCTGTGCAGC

Db   1801  TGGAGAAGGAGCCCATCGTCGGGCGGCCGGAGACCTTCTACGTGGACGGCGGCCAACCGG 1860
               ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Qy    181  CCGGGCCGCGAGCGACCGCAGGGCACCCTGGAGCTTCAACTTCCCCAGATCACCCTGTGCAGC   240

Qy   2018  AGACCAAGATCGGCAAGGCCGGCTACTGACCAGCCACCGGGGCGGGCCGGCGAGAGATCGTGAGCC 2077
               ||||||||||||||||||||||||||| |||||||| ||||||||||||||||||  ||||| |||||
Db    398  GCCCCGTGGTGAGCATCAAGGTGGCGGCGGATCAAGAGGCCCGTGCTGCTGGACACCGGGCG.   457

Db   1861  AGACCAAGCTGGGGAAGGGCGGCTACGTGACCGACGGCGTGGTGAGCAGGTGGTGAGCA 1920
               ||||||||  |||||| |||||||| ||||||||| |||||||||||||||||||||||
Qy    241  GCCCCCTGGTGACCATCAGGATCGGCGGCGGCCTCAAGGAGGGGCGGCTCGACACCCGGCG   300

Qy   2078  TGACCGAGACCACCAACCAGAAGACCGAGCTGCAGCAGCCAGTCTGCCCTGCCCTGCAGGACA 2137
               ||||||||||||||  ||||||||||| ||| |||| ||| ||||| |||||||||||||||
Db    458  CCGACGACACCGTGCTGGAGGAGATGAGCTGCCAGTGAGCTGGAAGCCAAGATGATCG   517

Db   1921  TCGCCGACACCACCACCAACCAGAAGACCGAGCTGCAGCAGCCAGTCCACCTGCCCTGCAGGACA 1980
               |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Qy    301  CCGACGACACCGTGCTGGAGGAGATGAACCTGCCCGGGCAAGTGGGAAGCCCAAGATGATCG   360

Qy   2138  GCGGCAGCCAGGTGAACATCGTGACCGACAGCAGCCAGTAGCCCTGGGCATCATCCAGGCCC 2197
               |||||||||||||||||||||||  |||||||| |||||||||||||||||||||||||||
Db    518  GCGGCATCGGCGGCTTCATCAAGTGCGCCAGTAGCCAGTACTGAGATCGGATCGGGAGATCTGCG   577

Db   1981  GCGGCCCTGGTGAACATCGTGACCGACAGCAGCCAGTAGCCCTGGGCATCATCCAGGCCC 2040
               ||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    361  GCGGGATCGGCGGCTTCATCAAGTGCGGCAGTAGCCAGTACTGAGATCGGGAGATCTGCG   420

Qy   2198  AGCCCGACAAGAGCGAGAGCGAGACCGTGGTGAACCAGATCATCGAGGCACAGGAGAAGG 2257
               |||||||||||||||||||||||||||||||||| ||||| |||||||||||||||||||
Db    578  GCAAGAGGCCATCGGCACCGACCGTGCTGATCGGGTGAACATCATCGGCGGCCGCGCA   637

Db   2041  AGCCCGACAAGAGCGAGAGCGAGACCGTGGTGAACCAGATCATCGAGGCACAGGAGAAGG 2100
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    421  GCCACAAGGCCATCGGCACCGACCGTGCTGTGGGTGGGTGAACATCATCGGCGGCCGCGCA   480

Qy   2258  AGAAGGTGTACCTGACCTGGTGGTGCCGCCACAAGGGCATCGGCGACCGAGCAGGATCG 2317
               ||||| ||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    638  ACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGC   697

Db   2101  AGAAGGTGTACCTGACCTGGCCTGGCCTGCCGCCCACAAAGGGCATCGGCGGCGAGCAGGTGG 2160
               ||||||||||||||||||  |||||| ||| ||||| ||||||||||||||||||||||
Qy    481  ACCTGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACGAGGTGC   540

Qy   2318  ACAAGCTGGTGACGAGAGCGCATCCGCAAGGTGCTGTTCCTGACACGGCATCGATGGCGGCA 2377
               ||||||||||| |||||||||||||||||||||||||||||||||||  ||||||||||||
Db    698  CCGTGAAGCTGAAGCCCCGGACGTGAGCGGCCAGTGAGCTGCCCCTGACCGAGG   757

Db   2161  ACAAGCTGGTGACGCCGCGGCATCCGCAAGGTGCTGTTCCTGACGCCCGATCGATGGCGCGCA 2220
               |||||||||||| |||| ||||||||||||||||||||||||||| |||||||||||||||
Qy    541  CCGTGAAGCTGAAGCCCGGGATGGACGGCCAGTGAGCTCAAGGTCAAGCGGTCAAGCCCCTGACCGAGG   600

Qy   2378  TCGTGATCGTCTACCAGTACTACATGACGACACCTTGTGACGCACCTGTGTACGTGGGCAGGATCGATT 2437
               |||||||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||||||||
Db    758  AGAAGATCAAGGCCCTGCACCGCCATCTGCGAGGACGAGGATGGGAGGGCAAGGATCACCA   817

Db   2221  TCGTGATCGTCTACCAGTACTACATGACGACGCTGTGTACGTGGGCAGGCGGCGGCCCTAGGATCGATT 2280
               ||||||||||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||||
Qy    601  AGAAGATCAAGGCCCCTGGTGGTGGATCGCACCGGAGATCGATGGGAGGGCAAGGATCAGCA   660

Qy   2438  AAAAGCTTCCCGGGCGGCTAGCACCGTGAATTC 2469
               ||||||||||||||||||||||||||||||||||||
Db    818  AGATCGGCCCCGAGAACCCCTACAACCCCGTGTTCGGCATCAAGAGCTGCTGAACAAGGCACCCAGGCACAGCA   877

Db   2281  AAAAGCTTCCCGGGCGGCTAGCACCGTGAATTC 2312
               |||||||||||||||||||||||||||||||||
Qy    661  AGATCGGCCCCGAGAACCCCTACAACCCCGTGTTCGGCATCAAGAGCTGCTGAACAAGGCACCCAGGCACAGCA   720
```

```
RESULT 5
AR373389
LOCUS       AR373389         2312 bp    DNA      linear      PAT 18-DEC-2003
DEFINITION  Sequence 84 from patent US 6602705.
ACCESSION   AR373389
VERSION     AR373389.1   GI:40075492
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2312)
  AUTHORS   Barnett,S.W., Megede,J., Greer,C. and Selby,M.
  TITLE     Expression of HIV polypeptides and production of virus-like
            particles
  JOURNAL   Patent: US 6602705-A 84 05-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..2312
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match     83.1%;  Score 2052;  DB 6;  Length 2312;
Best Local Similarity   93.6%;  Pred. No. 2.3e-206;
Matches 2165;  Conservative   0;  Mismatches 135;  Indels 12;  Gaps  2;
```

```
Qy    170  GCGGGCAAGGAGGAGGCCACCCAGATGAAGGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCGCG 229
               |||||||||||||||||||||||||||||||||||||||||
Db    938  AGGTGCAGCTGCGGCCATCCCCACCCCCGCCCGGCCTGAAGAGAAGAGAGGCGTGACCGTGC   997
```

```
Db  1081 TGGGCAGCGACCTGGAGATCGGCCAGCACGCCACCCAGAGATGCGAGCAGCCCACC 1140

Qy  1298 TGCTGCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGAAGCAGCCCCCTTCCTGT 1357
         ||||| |||| || ||||||||||| ||| |||||||| ||||||||||||||||||
Db  1141 TGCTGCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGAAGCAGCCCCCTTCCTGT 1200

Qy  1358 GGATGGCTACGACGTGCACCCCGACAAGTGACCGTGCAGCCCATCGAGCTGCCCGAGA 1417
         ||| ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  1201 GGATGGCTACGACGTGCACCCCGACAAGTGACCGTGCAGCCCATCATGCTGCCCGAGA 1260

Qy  1418 AGGAGAGCTGGACCGTGACGACATCCAGAAGCTGGTGGGCCAAGCTGAACTGGCCAGCC 1477
         |||| ||||||||||||| |||| |||||||||||||||| |||||||| |||||||||
Db  1261 AGGACAGCTGAACCGTGAACGACACATCCAGAAGCTGGTGGGCCAAGCTGAACTGGCCAGCC 1320

Qy  1478 AGATCTACCACCGGCATCAAGGTCGCGCAGCTGCGCGTGCAAGCTGCTGCGCGGGCCC 1537
         ||||| ||||||||||||||||| ||||| |||||||||||||||||||| |||||||
Db  1321 AGATCTACGCCGGCATCAAGGTGAAGCGAGTGCTGTGCGCGGGCACCAAGGCCC 1380

Qy  1538 TGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGCCGAGAACCGCGAGA 1597
         ||||||||||| |||||||||||| ||||||||||||| |||||||| ||||||||||
Db  1381 TGACCGAGGTGATCCCCCTGACCGAGGAGGCCGAGCTGGCTGCCGAGAACCGCGAGA 1440

Qy  1598 TCCTGCGAGCCGTGCACGCGGTGTACTACGACACCCCAGCCAAGGACCAAGGACTTGCCGAGA 1657
         ||||||||||| ||||||||||| ||||||||||||||||||||||||| ||||||||||||
Db  1441 TCCTGAAGGAGCCCGTGCACGAGGTGTACTACGACACCCCAGCCAAGGACCAAGGACCTGGCCGAGA 1500

Qy  1658 TCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGAATCCACCAGAATCC 1717
         |||||||||||||||||||||| |||||||||||||||||||||||| ||| ||||||||||
Db  1501 TCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGAATCCTTCAAGAACC 1560

Qy  1718 TGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCACACCGCCCACACCACCCAACGACGTGAAGCAGCTGA 1777
         ||||||||||||||||||| ||||||||||||||||||||||||| |||||||||||||||||||||||||
Db  1561 TGAAGACCGGCAAGTACGCCCCCATGCGCGGGCCCATGCGCCCCATGCGGGGGCCCACCAACGACGTGAAGCAGCAGCTGA 1620

Qy  1778 CCGAGGAGGCCGTGCAGAAGATCGCCAGAGAGATCGGCGAGAAGAATCGGCCAACTGGGGCCAAGACCCCCAAGT 1837
         |||||||||||||||||||||||| |||||||||||||| |||||||||||||| ||| ||||| |||| ||||||
Db  1621 CCGAGGAGGCCGTGCAGCACATCGGCGAGAAGCTGAACATCGTGGGCCAACTGGGGCCAAGATCCCCAAGT 1680

Qy  1838 TCCGCCTGCCCCATCCAGAAGGAGACTGGGAGACCTGGTGGACCGAGTACTGGCAGGCCA 1897
         || |||| ||||| |||||||||||||||||||||||| ||||||||||||||||||| ||
Db  1681 TCAAGCTGCCCATCCAGAAGGAGACTGGGAGGACCTGGTGGATGGAGTACTGGCAGGCCA 1740

Qy  1898 CCTGGATCCCCGAGTGGGAGTTCGTGTAACACCCCCCCCCTGGTGAAGCTGGTGGTACCAGC 1957
         |||||||||||||||||||||| ||||||| |||||||||||||||||||||||||||||||
Db  1741 CCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGGTGGTACCAGC 1800

Qy  1958 TGGAGAAGGAGCCCATCATCGTGTGGGCCGAGACCTTCTACGTGGACGGCGCCAACCGCG 2017
         |||||||||||||| |||||||||||||||||| ||||||||||||||||||||||||||
Db  1801 TGGAGAAGGAGCCCATCATCGTGTGGGCCGAGACCTTCTACGTGGACGGCGCCAACCGCG 1860

Qy  2018 AGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGCCGGCAGAAGTCGTGAGCC 2077
         ||||||||||||||| |||| ||||||||||||||||||||||||||||||| ||||||
Db  1861 AGACCAAGGTGGGGCAAGGCGGCTACGTGACCGACCGGGGCCGGCAGAAGGTGGTGAGCA 1920

Qy  2078 AGCCCGGAGTTCCCCGAGCTGCAGAGCCCTGGCCCGCAGGTCTGGCCCGCAGGACA 2137
         || |||||| ||||||||||||| |||||||||||||||||| |||||||||||||
Db  1921 TCGCCGACACCACCAAGGAGCTGCAGAGCCCTGGCCCGCCGCCGCACCCCTGCAGGACA 1980

Qy  2138 GCGGCAGCCGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC 2197
         ||||| |||| ||||||||||| ||||||||||||||||||||||||| |||||||||||||
Db  1981 GCGGCAGCCGGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC 2040

Qy  2198 AGCCCGACAAGAGCGAGAGCGAGCTGGTGAACATCATCAGCGTGGAGATCGAGAAGAGG 2257
         ||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||||
Db  2041 AGCCCGACACCACCAAGGAGCTGGTGAAATCGGCCAGCGAGCTGGTGAACATCCAGCAGAAGG 2100

Qy  2258 AGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGGCAACGAGCAGATCG 2317
         ||||||||||| ||| || |||||||||||||||||||| || ||||||||||||||||||
Db  2101 AGAAGGTGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGGTGG 2160

Qy  2318 ACAAGCTGGTGAGCGCCGGCATCCGCAAGGTGCTGTTCCTGACGACGATGGCGCA 2377
         |||||||||||||||| |||||||||| |||||||||||| ||||| ||||||||
Db  2161 ACAAGCTGGTGAGCGCCGGCATCCGCAAGGTGCTGTTCCTGACGATGGCGCA 2220
```

(Right column:)

```
Qy  2378 TCGTGATCTACCAGTACATGACGACGACCTGTACGTGCGGCAGCGCGGCCCTAGGATCGATT 2437
         ||||||||||||||||| ||||||||||||||||||||||||||| |||||||||||||||
Db  2221 TCGTGATCTACCAGTACATGACGACGACCTGTACGTGCGGCAGCGCGGCCCTAGGATCGATT 2280

Qy  2438 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2469
         |||||||||||||||||||||||||||||||
Db  2281 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2312


RESULT 6
AX427930
LOCUS           AX427930        9166 bp    DNA     linear    PAT 20-JUN-2002
DEFINITION      Sequence 168 from Patent WO0232943.
ACCESSION       AX427930
VERSION         AX427930.1   GI:21538017
KEYWORDS        .
SOURCE          synthetic construct
  ORGANISM      synthetic construct
                artificial sequences.
REFERENCE       1
  AUTHORS       Huang,Y. and Nabel,G.J.
  TITLE         Modifications of hiv env, gag, and pol enhance immunogenicity for
                genetic immunization
  JOURNAL       Patent: WO 0232943-A 168 25-APR-2002;
                GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..9166
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"

ORIGIN

Query Match             83.1%;  Score 2051.2;  DB 6;  Length 9166;
Best Local Similarity   92.6%;  Pred. No. 2e-206;
Matches 2192; Conservative    0;  Mismatches 163;   Indels   13;   Gaps   3;

Qy    14 TGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCCAACATCCTGATGCAGGCAGCAACTTCA 73
         |||||||||||||||||| |||||||||||||||||||||||||||||||||||| ||||||
Db  2961 TGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCCAACATCCTGATGCAGGCAGCAACTGCA 3020

Qy    74 AGGGCCCCAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGCCCGCCA 133
         |||||||||||||||||||||||| ||||||||||||||||||||||||| |||||||||
Db  3021 AGGGCCCCAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGCCCGCCA 3080

Qy   134 ACTGCCGCGCCCCCGCCCCCGCCAAGAAGGCTGCTGGAAGTGCGGCAAGAGGGCCACCAGATGA 193
         ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  3081 ACTGCCGCGCCCCCGCCCCCGCCAAGAAGGGCTGCTGGAAGTGCGGCAAGAGGGCCACCGGTGA 3140

Qy   194 AGGACTGCACCGAGCGCCAGGCCAGCAGGCTAA-TTTTTAGGGAAGATCTGCCTTCCACAAGGGA 253
         |||||||||||||||||||||||||||||||||   |||||||||||||||||| |||| ||||
Db  3141 AGGACTGCACCGAGCGCCAGGCCAGCAGGCTAA-TTTTTAGGGAAGATCTGCCTTCCACAAGGGA 3199

Qy   254 AGCCCGGAGTTCCCCGAGTTTCTTCAGCAGCAGCCGCCGCAGCCCCACCAGAGAGCTTC 313
         || |||||||||||||||| |||| |||| | | ||| ||||| |||||||||||| ||
Db  3200 AGGCCAGGAGGGAATTTCTTCAGCAGCAGCCGCCGCAGCCCCACCAGCCCCACCAGAGAGCTTC 3259

Qy   314 AGGTGCGCGG-------CGACAACCCCCAGATCACCCGGCAGGGGCCAGGCCAGGGCA--- 364
         ||||| |||          |||||| |||||||| || |||||||| |||||||||||
Db  3260 AGGTTTGGGGAAGAGAAACAACTCCTCTCAGAACACCTTGGCAGGAGCCGATAGACAAGAACTG 3319

Qy   365 ---CCCTGAACTTCCCCGATCACCCCCAGATCACCCGGCAGGGGCCCCCTGAGCATCATCGAGAAG 421
                |||| |||| || |||||||||| ||||| |||||||| |||||||||||||||||
Db  3320 TATCCTTTAGCTTCCCTGATCACTCTTTGGCAGCGACCCCGATCAAGGCATCAGGGAGA 3379

Qy   422 GCGGGCCAGATCAAGGCATCGGCGGGCATCGGCGGGCAACGAGCAGATCG 481
         |||| |||||||||||| |||||| |||| |||||||| ||| ||||||||
Db  3380 GGGGCCAGGCCAGTGAAGGGGCCAGTGCCCTTTCTAGACACCGAGCAGGTGG 3439

Qy   482 TGAACCTGCCCGGCCCGCTGAGCGAAGTGCTGTTCCTGACGACGATGGCGCA 541
         | ||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db  3440 TGAACCTGCCCGGCCCGCTGAGCGAAGTGCTGTTCCTGACGATGGCGCA 3499
```

```
Qy   1622 TGTACTACGACCCCAGCAAGAAGGACCTGGTCGCCGGAGATCAGAAGCAGGGCCACGACCAGT 1681
          |||||||||||||||||||  |||||||||||||||||||||| |||||||||||| ||||||
Db   4580 TGTACTACGACCCCAGCAAGAAGGACCTGATCGCCGGAGATCAGAAGCAGGGCCAGGGCCAGT 4639

Qy   1682 GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCAAGA 1741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db   4640 GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCCGCA 4699

Qy   1742 TGCGCACCGCCACCAACCAGCTGAAGCAGCTGAACGTGACCAGCCGTGCAGAAGATCGCCA 1801
          |||||||||||||||||||||| ||||||||||||||||||||||||||||| ||| ||||
Db   4700 TGAAGGGCGCCACCAACCAGCTGAAGCAGCTGAACGTGACCAGCCGTGCAGAAGATCGCCA 4759

Qy   1802 TGGGAGAGCATGTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCATCCAGAAGGAGA 1861
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4760 CCGGAGAGCATGTGATCTGGGCAAGACCCCCAAGTTCAAGCTGCCATCCAGAAGGAGA 4819

Qy   1862 CCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCG 1921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4820 CCTGGGAGGCCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCG 4879

Qy   1922 TGAACACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG 1981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4880 TGAACACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG 4939

Qy   1982 CCGAGACCTTCTACGTGACGGCGCGCGAGAGATCGGCAAGGCCCGCT 2041
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   4940 CCGAGACCTTCTACGTGACGGCGCGCGAGAGACCGAGAGATCGGCAAGGCCCGCT 4999

Qy   2042 ACGTGACCGAGACCGGCGGCGAGAAGATCGTGACCGAGCCTGACCGAGCCCAACCAAGAA 2101
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5000 ACGTGACCGAGACCGGCGGCGAGAAGAGTGGTGGTCCCCCGAGGCGGCGAGAAGGAGA 5059

Qy   2102 CCGAGCTGCAGCCGCATCCAGCTGGCCCTGCAGCGACGCGGCGAGAGGTGAACATCGTGA 2161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5060 CCGAGCTGCAGCCGCATCCAGCTGGCCCTGCAGGACAGCGCGGCCTGGAGGTGAACATCGTGA 5119

Qy   2162 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCCAGCAGAGCGAGAGCGAGC 2221
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5120 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCCAGCAGAGCGAGAGCGAGC 5179

Qy   2222 TGTGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCGTGGGTGC 2281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5180 TGGTGAGCCAGATCATCGAGCAGCTGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCGTGGGTGC 5239

Qy   2282 CCGCCCACAAGGGCATCGGCGAACAGCAGATCGCACAAGTGGTGAGCAAGGGCATCC 2341
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5240 CCGCCCACAAGGGCATCGGCGAACAGCAGATCGCACAAGTGGTGAGCAAGGGCATCC 5299

Qy   2342 GCAAGGTTGCTGTTCCTGGACGGCATCGA 2369
          ||||||||||||||||||||||||||||
Db   5300 GCAAGGTTGCTGTTCCTGGACGGGCATCGA 5327
```

```
Qy    542 TGCGCCAGTACGACCACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGCGCACCGTGC 601
          ||||||||||||||||||||||||| |  |||||||| ||||||||||||||||||||||||||
Db   3500 TGGGCCAGTACGACCACCAGATCCTGATCGAGATCTGCGGCACAAGGGCCACCGCACCGTGC 3559

Qy    602 TGATCGGCCCCACCCCGTGAACATCAGTCGGCCCGCCAACATGCTGCACCCAGTGGCTGCGCA 661
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3560 TGGTGGGCCCCACCCCGTGAACATCAGTCGGCCCGCCAACATGCTGCACCCAGATCGGCTGCGCA 3619

Qy    662 CCCTGAACTTCCCCATCAGCGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGG 721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3620 CCCTGAACTTCCCCATCAGCGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGG 3679

Qy    722 ACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCCA 781
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3680 ACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGGTGTGGAGA 3739

Qy    782 TCTGCACCGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGACCCTACA 841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3740 TCTGCACCGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACA 3799

Qy    842 ACACCCCCGTGTTCGCCATCAAGAAGAAGAGCACCAAGTGGCGCAAGCTGGTGGTGACT 901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3800 ACACCCCCGTGTTCGCCATCAAGAAGAAGAGCACCAAGTGGCGCAAGCTGGTGGTGACT 3859

Qy    902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC 961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3860 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC 3919

Qy    962 CCGCGGCCTGACGAAGAAGAGCCGTGACCGTGTCTGACGTGGGGCGACGCCTACTTCA 1021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3920 CCGGCGGCCTGACGAAGAAGAGCCGTGACCGTGTCTGACGTGGGGCGACGCCTACTTCA 3979

Qy   1022 GCGTGCCCCTGGACAGGGACTTCCGCAAGTACACACCGCCTTCACCATCCCCAGCATCAACA 1081
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3980 GCGTGCCCCTGGACAAGGACTTCCGCAAGTACACACCGCCTTCACCATCCCCAGCATCAACA 4039

Qy   1082 ACGAGACCCCCGGCATCCGTACCACAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCC 1141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4040 ACGAGACCCCCGGCATCCGTACCACAGTACAAGGTGCTGCCCCAGGGCTGGAAGGGCAGCC 4099

Qy   1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCCCAACCCCG 1201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4100 CCGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGAAGCAGAACCCCG 4159

Qy   1202 AGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCC 1261
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4160 ACATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCC 4219

Qy   1262 AGCACCGCGCCAAGATCGAGGAGCTGCGCGAACACCTGCTTGCGCTGGGCTTCACCACCC 1321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4220 AGCACCGCACCAAGATCGAGGAGCTGCGCGAACACCTGCTGCGCTGGGCTTCACCACCC 4279

Qy   1322 CCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGATGGGCTACGAGCTGCACCCCG 1381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4280 CCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGATGGGCTACGAGCTGCACCCCG 4339

Qy   1382 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACGTGACCGTGAACGACA 1441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4340 ACAAGTGGACCGTGCAGCCCATCGTGCTGCCCGAGAAGGACGTGACCGTGAACGACA 4399

Qy   1442 TCCAGAAGCTGGTGGCACTGTGGGCGAACTGACCCCAGGTCACCCCGGCATCAAGGTGC 1501
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4400 TCCAGAAGCTGGTGGGCACTGTGGGCGAACTGGGCCGCAGATCTACCGGGCATCAAGGTGC 4459

Qy   1502 GCCAGTGTGCGAAGTCTGCGCGGCGGCGCAAGCAGCCCTGACCGGCGACATCGTGCCCGACCG 1561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4460 GCCAGTGTGCAAGTCTGCGCGGCGCGCCAACCACCAAGGCCCTGACCGGTGTGCCCGACCG 4519

Qy   1562 AGGAGGCCGAGCTGGAGCTGGCGGAGATCCTGCGCAGCAGATCCTGCGCACGCGGCG 1621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4520 AGGAGGCCGAGCTGGAGCTGGCGGAGATCCTGCGAGATCCTGAAGGAGAACAACGCCGTGCACGCGGCG 4579
```

RESULT 7
AX427936

```
        /db_xref="taxon:32630"
        /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match    82.7%;    Score 2042.8;  DB 6;  Length 9788;
Best Local Similarity 91.8%;  Pred. No. 1.5e-205;
Matches 2197;  Conservative 0;  Mismatches 177;  Indels 18;  Gaps 3;

Db   3915 CCGCGGCCTGAAGCAGAGAAGAGAGGCGTGCCAAGTACACCGCCTGTCTGACGTGGGCGACGCCTACTTCA 3974
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1022 CCGCGGCCTGAAGCAGAGAAGAGAGGCGTGCCAAGTACACCGCCTGTCTGACGTGGGCGACGCCTACTTCA 1081

Db   3975 GCGTGCCCCTGGACAAGGACTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACA 4034
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Qy   1082 GCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCAGCATCAACA 1141

Db   1082 ACGAGACCCCGGCATCCGTACCACGTCACGTGGTGCCCCAGGGCTGGAAGGGCAGCC 4094
          |||||||||| | |||||||| |  | | |||||||||||||||||||||||||||||
Qy   4035 ACGAGACCCCGGCATCCGTACCACGTCACGTGGTGCCCCAGGGCTGGAAGGGCAGCC 1201

Db   2961 TGGCCGAGGCCATGAGCCAGGCCACAGCGCCAACTCCTGATGCAGCGCGAGCAACTTCA 73
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Qy   1142 CCAGCATCTTCCAGAGCAGCAGCATGACCAAGATCTCGGAGCCTTCCGGCCGCCCACCCG 3020

Db     74 AGGGCCCCAAGGCCACCATCAAGTGCTTCAAGTGTGGCCAAGAGGGCCACATGGCC 4154
          ||||| ||||| ||||| |||| ||||| ||||| ||||||||||||| |||||||
Qy   4095 CCGGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCTTCCGGCCGCCCACCCG 3080

Db   3021 AGGGCCCCAAGGCCACCATCAAGTGCTTCAAGTGCGGCAAGAGGGCCACCAGATGA 4214
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3081 ACTCCGCGCCCCGGCCCCGAGGCCCGGAGCCCAACAGCCGCGCCCACCAGATGA 1321

Db    134 ACTCCGCGCCCCGGCCCCGAGGCCCGGAGCCCAACAGCCGCGCCCACCCACCAGTGA 4274
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1262 AGGACTGCACCGAGCCCAGGCGACAGCAGGCTACGAGCTTCACCAGCTGCACCCCG 4334

Db   3141 AGGACTGCACCGAGCCCAGGCGACAGCAGGCTAA------TAGGGAAGATCTGGCCTTCCCACAAGGGAA 3194
          |||||||||||||||||||||||||||||||||       ||||||||||||||||||||||||||||||
Qy   1322 CCGACAAGAAGCACCAGAGAAGCACCCCCCTTCCTGTGGCCCAACAGCCGCGCCCACCCACCAGTGA 1381

Db    254 AGGGCCCGGAGTTCCCGAGCAGCAGACCGCGGCCCAACAGCCGCGAGCGGTGC 4334
          ||||||||||||||||||||||| |||||||||||| |||||||||||||||||
Qy   4275 CCGACAAGAAGCACCAGAGAAGCACCCCCCTTCCTGTGGCCCAACAGCCGCGACA 1441

Db   3195 AGGGCCAGGGAATTTTCTTCAGACAGACAGACAGCCAACAGCCACCAGAGAAGCAGCTTC 3254
          ||||||||||| | |||||||||||||||| |||||||| ||||||||||||||||||||
Qy   1382 ACAAGTGACCGTGCAGCCCATCGAGCTGCCGAGAAGCAGCTGAACAGCGTGAACGACA 4394

Db    314 AGTTGCGCGG------CGACAACCCCGCAGCGAGGCGGCCCAGGGCGCCAGGGCA--- 364
          ||||||||         |||||||||||||||||||||| ||||||| |||||||
Qy   4335 ACAAGTGACCGTGCAGCCCATCGAGCTGCCGAGAAGCAGCTGAACAGCGTGAACGACA 1501

Db   3255 AGGTTTGGGGAAGAGACAACAACTCCCTTCTTTGGGCAGACCGGATAGACAAGGAACTG 3314
          ||| |||||||||| ||| ||||||| ||| |||| |||| ||||| ||||||||||
Qy   1442 TCCAGAGAGCTGTGTCGCGCAAGCTGGTGGGCAAGCTGTGTGCGCGGCATCAAGGTGC 4454

Db    365 ---CCCTGAACTTCACCCTGTGCGCAGATCACCCTGCTGGCGCAGCAGACACCGTGTCTGAGCATCAAGGTGG 421
             | | |||||||||||||||| ||||| | |||| || ||| ||| ||| ||| ||||||| ||||||||
Qy   4395 TCCAGAGAGCTGTGTCGCGCAAGCTGGTGGGCAAGCTGTGTGCGCGGCATCAAGGTGC 1561

Db   3315 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGGCAGCAGGCCGCGCGGCCCTGACCGTCCGAGGTGGTGCCCTGACCG 3374
          |||| | | |||||| ||||||| || ||||| |||||||||| |||||||||||||||||||||||| ||||||||||
Qy   1502 GCCAGCTGTGAAGCTGTGTGCGCGGCACCAGGAACCCGCGAGATCCTGCGCGGCGCGGCG 4514

Db    422 GCGGCCAGATCAAGGAGGGCCCTGCTGGAAGCCCAAGATGATCGGCCGCGGCATCGGCGTGGCGAGCCCGTCGACGGCG 481
          ||| |||||||||| |||||| | ||||||||||| ||| |||||||||||||||||||| ||| ||||||||||||
Qy   1562 GCCAGCTGTGAAGCTGTGTGCGCGGCACCAGGAACCCGCGAGATCCTGCGCGGCGCGGCG 1621

Db   3375 GGGGCCAGTCAAGGAGGGCCCTGCTGGAAGCCCAAGATGATCGGCCGCGGCATCGGCGTGGCGAGCCCGTCGACGGAGA 3434
          ||||| |||||||||||||||||||||||||||||||| ||||||||||||||||||||| ||| ||||||||||||
Qy   4515 AGGAGGCCGAGCTGGACGCTGGGCGAGCTGTGGCGAGCCCGTGCACGGCG 4574

Db    482 TGAGCTGCCCGGCAAGTGAAGCCCAAGATGATCGGCCGCGGCATCGGCGCTTCATCAAGG 541
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1622 TGTACTACGACCCCAGACAGCTGGTGGCCCAGAGACCTGGAGGCGGCCACCAGT 1681

Db   3435 TGAACTGCCCGGCAAGTGAAGCCCAAGATGATCGGCCGCGGCATCGGCGCTTCATCAAGG 3494
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   4575 TGTACTACGACCCCAGACAGCTGGTGGCCCAGAGACCTGGAGGCGGCCACCAGT 4634

Db    542 TGGCCAGTACGACCAGATCCTGATCGAGATCTGGCGCAAGAAGCCATCGGCCACCGTGC 601
          ||||| |||||||||| ||||| ||||||||| |||||| |||||||||||||||||||
Qy   3495 TGGGCCAGTACGACCAGATCCTGATCGAGATCTGGCGCAAGAAGCCATCGGCCACCGTGC 3554

Db    602 TGATCGGCCCACCCCGTGAACATCAGCCCGTCGCCGCCCAGCTGCTGACCCAGCATGTCTGGCTGCA 661
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Qy    662 GGACCTACCAGGACCCCCACCCCCGTGAACATCAGCGACGCTGAAGCTGAAGCAGCCCGTCGACATGG 721

Db   3555 TGGTGGGCCCACCCCGTGAACATCAGCCCGTCGCCGCCCAGCTGCTGACCCAGCATGTCTGGCTGCA 3614
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Qy   3615 GGACCTACCAGGACCCCCACCCCCGTGAACATCAGCGACGCTGAAGCTGAAGCAGCCCGTCGACATGG 3674

Db    662 CCCTGAACTTCCCGAGCAGCAGCATGACCATCAGCCCCTGCCCGCTGCCCGTGAAGCTGAAGCCCGTCGAGA 721
          ||||||||||||||||| ||||||||||||||||||||| || || ||||||||||||||||||||||||
Qy    722 ACGGCCCCAAGGTGAAGCAGTGGCCGCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATC 781

Db   3615 CCCTGAACTTCCCGAGCAGCAGCATGACCATCAGCCCCTGCCCGCTGCCCGTGAAGCTGAAGCCCGTCGGTGGAGA 3674
          ||||||||||||||||| ||||||||||||||||||||| || || ||||||||||||||||||||||||||||
Qy    782 ACGGCCCCAAGGTGAAGCAGTGGCCGCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGATC 1921

Db    722 ACGGCCCCAAGGTGAAGCAGTGGCCGCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGA 781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3675 CCTGGGACCTGTGGACCGAGATCCCCGGATCCCCGGAGTCCCGAGTTCG 3734

Db    842 ACACCCCGTGTTCGCCATCAGAGAAGAGGACACCAGAGAAGAAGATCCACAAGACCAACGTGGTGACT 901
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy   3735 CCTGGGAGGCCCGTGGTGGAGGCCCACGTGTACACCGTGAACATCAGCAGGTGGGGCCATCATCGGCGG 3794

Db   3795 ACACCCCGTGTTCGCCATCAGAGAAGAGGACACCAGAGAAGAAGATCCACAAGACCAACGTGGTGACT 3854
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy   1982 CCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC 2041

Db    902 TCCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC 961
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   4935 CCGAGACCTTCTACGTGACGAGGCGCGGCGCCACCCACCGCGCGGCATCCCCACC 3914

Db   3855 TCCGCGGCCTGAAGCAGAGAAGAGAGGCGTGCCAAGTACACCGCCTGTCTGACGTGGGCGACGCCTACTTCA 3914
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2042 AGTGACCGACCGACCGGGCCCCGAGAGATCGTGCTGCTGACCGACGACCAACCAGATCG 2101

Db    962 CCGCGGCCTGAAGCAGAGAAGAGAGGCGTGCCAAGTACACCGCCTGTCTGACGTGGGCGACGCCTACTTCA 1021
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 3263 AGGTTTGGGGAAGAGACAACAACTCCCTCAGAAGCAGGAGCCGATAGACAAGGAACTG 3322

Qy  365 ---CCCTGAACTTCCCCAGATCACC              CTGTGAGCGCCCCTTGGTGAGCATCAAGGTGG 421
            | |||||||||||||||||||||              |||||||||||||||| |||||||||||||||
Db 3323 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAAGATAG 3382

Qy  422 GCGGCCAGATCAAGGAGCGGCCTGCTGGCATCATCCAGGCCCGACACACGAGCGAGC 481
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3383 GGGGCCAGCTCAGCAGCCATCATCCAGGCCCGACACACCGGCCATCGCTGCTGGAGGA 3442

Qy  482 TGAGCGGCTGCCGGCAAGTGAAGTGGACAAGAGAAGTGTACCTGAGCTGGTGC 541
            ||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db 3443 TGAACCTGCCCGGCGCCTGAAGCCCAGGCCTTCATCGCTGCTGGTGTCATCAAGG 3502

Qy  542 TGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGAGAGCGGCAGTCGACAAGGGCATC 601
            |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db 3503 TGGGCCAGTACGACCAGATCCTGATCGAGATCTGCGGGACCAGTCGCTGGTCGCACCGGTGC 3562

Qy  602 TGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTACCAGTA 661
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3563 TGGTGGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGCA 3622

Qy  662 CCCTGAACTTCCCCATCAGCCCCATCAGCCCCTGACGACCGTGCCCGAAGCTGAAGCCATGG 721
            |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db 3623 CCCTGAACTTCCCCATCAGCCCCATCAGCCCCTGACGACCGTGCCCGTGGAAGCTGAAGCCATGG 3682

Qy  722 ACGGGCCCCAAGGTGAAGCACTGCGCCCTGGCCGAGGAGAGATCAAGGCCCTGACCGGCCA 781
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3683 ACGGGCCCCAAGGTGAAGCAGTGCGCCCTGGCCGAGGAGAGATCAAGGCCCTGGTGGGAGA 3742

Qy  782 TCTGCGAGGAGATGGAGAAGGAGGGCCAAGATCAACAAGATCGGCCGCGAGAACCCCTACA 841
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3743 TCTGCACCGAGGAGATGGAGAAGGAGGGCCAAGATCAGCAGGATCGGCCGCGAGAACCCCTACA 3802

Qy  842 ACACCCCCGTGTTCGCCATCAAGAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT 901
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3803 ACACCCCCGTGTTCGCCATCAAGAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT 3862

Qy  902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC 961
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3863 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC 3922

Qy  962 CCGCCCGGCCTGAAGAAGAAGAAGGCCTGACCGTGCTGGACGTGGGCGACGCCTACTTCA 1021
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3923 CCGCCCGGCCTGAAGAAGAAGAAGGAGCGTGCTGGACGTGGGCGACGCCTACTTCA 3982

Qy 1022 GCGTGCCCCTGGACGAGGACTTCCGCAAGTACACGGCCTTCACCATCCCCAGCATCAACA 1081
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3983 GCGTGCCCCTGGACGAGGACTTCCGCAAGTACACACCGGCCTTCACCATCCCCAGCATCAACA 4042

Qy 1082 ACGAGACCCCGGCATCGCTACCAGTACAACGTGCTGCCCGAGGCTGGAAGGGCAGCC 1141
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4043 ACGAGACCCCGCATCGCTACCAGTACAACGTGCTGCCCGAGGCTGGAAGGGCAGCC 4102

Qy 1142 CCAGCATCTTCAGAGCATCTGCGGCAAGAGGGCCACATCGCCCCCA 1201
            |||||||||||||||||||||||||||||||||||||||||||||
Db 4103 CCGCCATCTTCCAGTGCATCTGCGGCAAGAGGGCCACATCGCCCCCG 4162

Qy 1202 AGATCGTGATCTACCAGTACATTGAGCAGTGGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 1261
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4163 ACATCGTGATCTACCAGTACATTGAGCAGTGGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 4222

Qy 1262 AGCACCGGCGCCAAGATGAGAGCTGGCAGCACCTTCGCTGCGGCTTCACCACCC 1321
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4223 AGCACCGGCGCCAAGATGAGAGCTGCGCCAGCACCTTCGCTGCGGCTTCACCACCC 4282

Qy 1322 CCGACAAGAGCACCACAGAGAACCGGCAAACGCCCCCTTCGTGTGAATGGCGCGAGCTGC 1381
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4283 CCGACAAGAGCACCACAGAGAACCGGCAAACGCCCCCTTCGTGTGAATGGCGCGAGCTGC 4342

Qy 1382 ACAAGTGACCGTGCAGCCCATCGAGTGCCCGAGAGAGAAGGAGCTGACCACGTGAACGACA 1441
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4343 ACAAGTGACCGTGCAGCCCATCGTGTGCCCGAGAGAGAAGGAGCTGACCACGTGAACGACA 4402
```

```
Db 4995 ACGTGACCGACCGCGGCGCCGCAGAAGGTGGTGCCCCTGACCGACCACCAACCAGAAGA 5054

Qy 2102 CCGAGCTGCAGGCCCATGCAGGCCATCCAGTCTGGCCTGCAGGCGGCCAGCGAGCATCATCGTGA 2161
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5055 CCGAGCTGCAGGCCCATGCAGGCCATCCAGTCCACCTGGCCCTGCAGGCGGCCAGCGAGCATCATCGTGA 5114

Qy 2162 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCTGCCCGAGACAAGAGCGGAGC 2221
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5115 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCCCAGGCCGTGCCGCACAGGAGGAGAGC 5174

Qy 2222 TGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGAAGAAGTGTACCTGAGCTGGTGC 2281
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5175 TGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGAAGAAGTGTACCTGGCCTGGTGC 5234

Qy 2282 CCGCCCCAAGGGCATCGGCGGCAACAGCAGCAGAGTCGACAAGGGCATC 2341
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db 5235 CCGCCCCACAAGGGCATCGGCGGCAACAGCAGCAGAGTGCAGTCGCACAGGGCATC 5294

Qy 2342 GCAAGGTGCTGTTCCTGACGACGGCATCGTCCTGATGGCGCATCCTGTGATCTACCAGTA 2393
            ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db 5295 GCAAGGTGCTGTTCCTGACGACGGCATCGTGACAAGCCCCCAGGAGAGCACGAGAA 5346
```

RESULT 8
AX427931

```
LOCUS       AX427931                9169 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 169 from Patent WO0232943.
ACCESSION   AX427931
VERSION     AX427931.1  GI:21538018
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel.G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 169 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US);
FEATURES             Location/Qualifiers
     source          1..9169
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVRl012x/s containing HIV genes"
ORIGIN
```

```
Query Match    82.2%;  Score 2028.8;  DB 6;  Length 9169;
Best Local Similarity   91.7%;  Pred. No. 4.4e-204;
Matches 2194;  Conservative  0;  Mismatches 182;  Indels 16;  Gaps  4;
```

```
Qy   14 TGGCCGAGGCCATGAGCAGCCAGCACCAGGCCACCATCCTGATGCAGCGGCAGCAACTTCA 73
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2967 TGGCCGAGGCCATGAGCAGCCAGCAGGTCGAGCACCAGGCCACCATCATGCAGCGGCAGCAACTTCA 3026

Qy   74 AGGGCCCCAAGGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 133
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3027 AGGGCCCCAAGGCCGCAT---CAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGCCCGCA 3083

Qy  134 ACTGCCGCGCCCCGCCCCGCTGCGCAAGAAGGCGGCCAAGAAGAGGTGCGGCAAGACCTGGCC 193
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3084 ACTGCCGCGCCCCTGCGCAAGAAGGCGGCCAAGAAGGTGCGGCAAGACCTGGCC 3143

Qy  194 AGGACTGCACCGAGCAGGCTAA-TTTTTAGGGAAGATCTGGCCTTCCCACAAGGGA 253
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db 3144 AGGACTGCACCGAGCAGGCTCAGGAAGGTCTGGCCTTCCCACCAAGGGGCA 3202

Qy  254 AGGGCCCGAGTTCCCCGAGCAGAGAACCGGCAACGCCCCAGCCGCGAGCTGC 313
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3203 AGGGCCAGGAATTTCTTCAGAGCAGAGACCAGAGCCAACAGCCCCAGAGAGAGAGCTTC 3262

Qy  314 AGGTGCGGG-----CGACAACCCCGGCGAGGCCGCGGCGCCAGGGCA--- 364
            |||||          |||||||||||||| || ||| |||
```

```
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 163 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES    Location/Qualifiers
     source 1..9194
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN
```

```
Query Match          82.1%;  Score 2027;  DB 6;  Length 9194;
Best Local Similarity 91.6%;  Pred. No. 6.8e-204;
Matches 2194;  Conservative  0;  Mismatches 185;  Indels 16;  Gaps  4;

Qy       14 TGGCCGAGCCATGAGCCAGG---CCACCAGCGCCAACATCCTGATGCAGGGCAGCAACT   70
            ||||||||||||||||| |||   | |||||||||||||| || | |||||||||||||
Db     2967 TGGCCGAGGCCATGAGCCAGGTGGACCAACAGCGCCACCATCATGATGCAGCGGCAACT 3026

Qy       71 TCAAGGGCCCCAAGCGCATCATCAAGTTCCGCCTGCAACTGGGCGAAGGGCCACCAGA  130
            | ||||||||||| |||||||||||||||| |||||||||||||||||| |||||||||
Db     3027 TCCGCAACCAGCGCATCATCAAGTTCGAAGTGCTTCAACTGCGGCAAGGAGGGCCACCAGA 3086

Qy      131 GCAACTGCGCGCCCCCCGGCAGGGCCAAGAAGGGCTGCGGCAAGGAGGGCCACCAGA  190
            |||||| |||||||||| ||||||||||| |||||||||||||||||||||||||||||
Db     3087 GCAACTGCCGCGCCCCCCGGCAGGGCCAAGAAGGGCTGCGGCAAGGAGGGCCACCAGA 3146

Qy      191 TGAAGGACTGCACCGAGCGGCCAAGCCACGGCCAGCGCCAAGCCCCCCAGG  250
            ||||||||||||||||| ||||| |||||||| |||||||||||||||||||
Db     3147 TGAAGGACTGCACCGAGCGGCGGAATTTTTTAGGGAGCTAA-TTTTTAGGGACAGATCTGGCCTTCCCACAAG 3205

Qy      251 GCAAGGCCCGCGAGTTCCCAGCGAGCAGAGAACCGCGCAACAGCCCAGCCGCGGAGC  310
            |||||||||| |||||||||||||||| |||||||||||||||||||| |||||||||
Db     3206 GGAAGGCCAGGGATTTTCTTCAGAGCAGGACCAGAGCAGACCCCACCAGAGAGAGAGC 3265

Qy      311 TGCAGGTGCGCGG----CGACAACCCCGCAGCGAGGCCCGGCGCGGCCCCAGGGCA  364
            |||||||||| |     |||||||| | ||||||| | ||||||||||||||||
Db     3266 TTCAGGTTTGGGAAGAGAACAACAACTCCCTCCAGAAGCGGCAGCGGCCGGATAGACAACGGAA 3325

Qy      365 -----CCCTGAACTTCCCCAGATCACCCTGCGGCAGGACATCGTGA  418
                 ||||||||||||||||||||||| |||||| ||||||||||
Db     3326 CTGTATCCTTTAGTTCCCTCAGATCACTCTTTGGGCAGCGGCCCTCGTCACAATAAAGA 3385

Qy      419 TGGGGGGCCAGATCAGGAGGCCCGTGCTGGACACCGGCGCCGCCACCGTGCTGGAGG  478
            |||||| | |||||||||| |||||||||||||||||| |||||||||||||||||||
Db     3386 TAGGGGGCCAGTCAGGAGGCCCTTCAGACGACACCGGCGCCGCCACCGTGCTGGAGG 3445

Qy      479 AGATGAGCCTGCCGGAAGATCCTGATCGGAAGATCGACTGCCTTCATCA  538
            |||| |||||||||||||||||| |||||| |||||| ||||||||||||
Db     3446 AGATGAACCTGCCCGGGAAGATCCTGATCGGAAGATCGACTGCCTTCATCA 3505

Qy      539 AGGTGCGCCAGTACGACCAGATCCTGATCGAGAGGGCCATGGCCACCG  598
            |||||||||||| ||||||||||||||||||||| | ||||||||||||
Db     3506 AGGTGGGCCAGTACGACCAGATCCTGATCGAGAGTCGCAAGCCCACCACCG 3565

Qy      599 TGCTGATCGGCCCGACCGCCATCGAACATCGGCCGCAACTGCTGACCCAGCTGGGCT  658
            ||||| |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     3566 TGCTGGTGGCCCACCCCGTGCCCGAGCCGTGCCCCGAAGCTGAAGCCCGGCA 3625

Qy      659 GCACCCTGAACTTCCCCATCAGCCCCATCAGCCCCGAGACCGTGCCCGGCGCA  718
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3626 GCACCCTGAACTTCCCCATCAGCCCCATCAGCCCCGAGACCGTGCCCGGCGCA 3685

Qy      719 TGGACGGCGCCCAAGTGAAGCAGCTGGCCCCTGACCGGAGGAAGATCAAGGCCCTGACCG  778
            |||||||||||||| |||| |||||||||||||||||||||||||| ||||||||||||||
Db     3686 TGGACGGCGCCCAAGTGAAGCAGCTGGCCCCTGACCGGAGGAAGATCAAGGCCCTGTGTGG 3745
```

```
Qy     1442 TCCAGAAGCTGGTGGCCAAGCTGGCCAGCAGATCTACCCCGGCATCAAGGTGC 1501
            ||||||||||||||| |||||||||||||| ||||| |||||| ||||||||||
Db     4403 TCCAGAAGCTGTGTGGCCAAGCTGGCCAGCAGATCTACGCCCGGCATCAAGGTGC 4462

Qy     1502 GCCAGCTGTGCAAGTGCTCGCGCGCCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCG 1561
            |||||||   |||| |||||||||||||| |||||||||||||||||||||||||||||||
Db     4463 GCCAGCTGCAAGTGCTGCGCGGCACCAAGGGTGTGCCCCTGACCG 4522

Qy     1562 AGGAGGCCGGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGAGCCCGTGCACGGCG 1621
            ||||||| |||||||||||||||||||||| |||||||||||||||||||| ||||||||||
Db     4523 AGGAGGCCGGAGCTGGAGCTGGCCGAGAATCCGAGATCCTGAAGGAGCCCGTGCACGGCG 4582

Qy     1622 TGTACTACGACCCCAGCAAGGACCTGGTGGCCGAAGCAGGCCACGACCAGT 1681
            |||||||||| ||||||| |||||||| || |||||| ||||||||||||||
Db     4583 TGTACTACGACCCCAGCAAGGACCTGGTGGCCGAAGCAGGCCACGACCAGT 4642

Qy     1682 GGACCTACCAGATCTACCAGGAGCCTTCAAGAACGTGAAGACCTGGACTACGCCAAGA 1741
            |||||||||||||||||||||||| |||||||||||||||| ||||||||||||||||
Db     4643 GGACCTACCAGATCTACCAGGAGCCTTCAAGAACGTGAAGACCTACGCCGCCA 4702

Qy     1742 TGCGGACCGGCCCAACAACGACGTGAAGCAGTGACCTGAACAGGCCGTGCAGAAGATCGCCA 1801
            ||||||||| |||||||| ||||||| |||||||||||| || |||||||||||||| ||||
Db     4703 TGAAGGGCGCCCACAACAACGACGTGAAGCAGTGGTGCCCGCAGGCCGTGCAGAAGATCGCCA 4762

Qy     1802 TGGGAGAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCATCCAGAAGGAGA 1861
            ||| ||||||||||||||||||||||| |||||| |||||||||||| |||||||||||
Db     4763 TGGGAGAGCATCGTGATCTGGGCAAGACCCCCAAGTTCAAGCTGCCCATCCAGAAGGAGA 4822

Qy     1862 CCTGGGAGACCTGGTGGACCGACTACTGGCCAGCCACCTGGATCCCCGAGTTGGGAGTTCG 1921
            ||| |||||| ||| ||||||||||||| ||||||||||| ||| ||||||||||||||||
Db     4823 CCTGGGAGGCCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCG 4882

Qy     1922 TGAACACCCCCCCCGGTGAAGCGTGTGGTACCAGCGGGAGGAGGAGCCCATCATCGGCG 1981
            |||| |||||||||| |||||||| ||||||| |||||||||||||||||||||||||||
Db     4883 TGAACACCCCCCCCGGGTGAAGCTGTGTGGTACCAGCGGGAGGAGGAGCCCATCATCGGCG 4942

Qy     1982 CCGAGACCTTCTACGTGGACGGCGCCGGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCT 2041
            |||||||||||||| ||||||||||| |||||| |||||||||||||||||||||||||||
Db     4943 CCGAGACCTTCTACGTGGACGGCGCCGGCCAAGCCGAGACCAAGGTCGGCAAGGCCGGCT 5002

Qy     2042 ACGTGACCGACCGGGGCCGCCAGAAGATCGTGACCCTGACCGAGACCACCAACCAGAAGAGA 2101
            ||||||||||||| |||| ||| |||||||||| |||||||||||| |||||||||||||||
Db     5003 ACGTGACCGACCGGGGCCGCCAGAAGGTGGTGCCCCTGACCGAGACCACCAACCAGAAGAGA 5062

Qy     2102 CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCGCGAGGTGAACATCGTGA 2161
            ||||||||||||||||||||||||| ||||| |||||||||||| ||||||||||||||||
Db     5063 CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCGTGA 5122

Qy     2162 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCGACAAGAAGGAGAGCGAGC 2221
            ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     5123 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCGACAAGAAGGAGAGCGAGC 5182

Qy     2222 TGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGGTGTACCTGAGCTGGGTGC 2281
            ||||||||||||| |||||||||||||||||||||||||||||| |||||| ||||
Db     5183 TGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCCTGGGTGC 5242

Qy     2282 CCGCCCACAAGGGCATCGGCGGCAACGGCGCCAAGGGCATCC 2341
            ||| |||||||||||||||||||| |||||||||||||||| |
Db     5243 CCGCCCACAAGGGCATCGGCGGCAACGGCGCCAAGGGCCATCC 5302

Qy     2342 GCAAGGTGCTGTTCCTGGACGGCATCGGCGACGTCGTGATCTACCAGTA 2393
            |||||||||||||| ||||||||||||||||||||||| || |||||||||
Db     5303 GCAAGGTGCTGTTCCTGGACGGCATCGGCGACGTCGTGATCTACCAGTA 5354
```

```
RESULT 9
AX427925
LOCUS       AX427925       9194 bp    DNA      linear     PAT 20-JUN-2002
DEFINITION  Sequence 163 from Patent WO0232943.
ACCESSION   AX427925
VERSION     AX427925.1  GI:21538012
KEYWORDS
```

```
Qy    779 CCATCTGCGAGGAGGATGTGGAGAAGGAGGGCAAGATCACCAGGATCGGCCCCGAGAACCCT 838
          |||||||||||||     ||| ||||||||||||||||||   ||||||||||||||||||||
Db   3746 AGATCTGCACCGAGATGAGGAGGAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCT 3805

Qy    839 ACAACACCCCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGG 898
          ||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db   3806 ACAACACCCCCGTGTTCGCCATCAAGAAGGACGACCACCAAGTGGCGCAAGTGGTCG 3865

Qy    899 ACTTCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTTGGGCATCCCCC 958
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3866 ACTTCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTTGGGCATCCCCC 3925

Qy    959 ACCCCGCGGCCTGAAGAAGAGAGCGTGACCGTGCTGGACGTGGCGACGGCCTACT 1018
          ||||||||||||  |||||||||||||| |||||||||||||||||||||||||||
Db   3926 ACCCCCGCGGCCTGGCCTGAAGAAGAGAGCGTGACGTGCTGGACGTGGCGACGCCTACT 3985

Qy   1019 TCAGCGTGCCCCTGGACGAGGACTTCCGCAGTACACCGGCCTTCACCATCCCCAGCATCA 1078
          ||||||||||||||||||||||||||||||    |||||||||||||||||||||||||||
Db   3986 TCAGCGTGCCCCTGGACAAGGACTTCGGCCAGTACACCGGCCTTCACCATCCCCAGCATCA 4045

Qy   1079 ACAACGAGACCCCCGGCATCCGCTACCAGTACACAGTGCTGCCCCAGGGTCTGGAAGGCA 1138
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db   4046 ACAACGAGACCCCCGGCATCCGCTACCAGTACACAGTGCTGCCCCAGGGTCTGGAAGGGCA 4105

Qy   1139 GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCTTCCCGCGCCCGCAACC 1198
          |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db   4106 GCCCCGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCCGGCTGAAGAACC 4165

Qy   1199 CCGAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCG 1258
          ||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db   4166 CCGAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCG 4225

Qy   1259 GCCAGCACCGCACCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCGTGGGGCTTCACCA 1318
          ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db   4226 GCCAGCACCGCACCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCGTGGGGCTTCACCA 4285

Qy   1319 CCCCCGACAAGAGCACCAGGAAGAGGAGCCCCCTTCCTGTGATGGGCTACGAGCTGCACC 1378
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4286 CCCCCGACAAGAGCACCAGGAAGAGGAGCCCCCCTTCCTGTGATGGGCTACGAGCTGCACC 4345

Qy   1379 CCGACAAGTTGACCGTGCAGCGCCATCGAGCTGCGCCGAGAGGAGAGCTGACCGTGACG 1438
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4346 CCGACAAGTTGACCGTGCAGCGCCATCGTGCCCCGGAGAAGGACAGCTGACCGTGACG 4405

Qy   1439 ACATCCAGAAGCTGGTGGGCAAGTGAACTGGGCCAGCCAGATCTACCCGGATCAAGG 1498
          ||||||||||||||||||||||||||||||||||||||||||||||  ||||||||||
Db   4406 ACATCCAGAAGCTGGTGGGCAAGTGAACTGGGCCAGCCAGATC  ACGCCGGCATCAAGG 4465

Qy   1499 TGCGCCAGCTGCGCAAGCTGCTGCGCGGGCGCCAAGGCCCCGACCGACACATCGTGCCCCTGA 1558
          |||||||||||||||||||||    |||||||||||||||||||||||||||||||||||||
Db   4466 TGCGCCAGCTGCGCAAGCTGCTGCGCGGGCACCAAGGCGTGTGCCCCTGA 4525

Qy   1559 CCGAGGAGGCCCAGCTGGAGCTGGCCGAGAATCCTGCGCGAGCCCGTGCACG 1618
          |||||||||||||||||||||||||||||||||||||  |||||||||||||
Db   4526 CCGAGGAGGCCCAGCTGGAGCTGGCCGAGAACCGGCCGAGAACCCGTGCACG 4585

Qy   1619 GCGTGTACTACGACCCCCAGCAAGCAAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACC 1678
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4586 GGGTGTACTACGACCCCCAGCAAGCAAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGGCC 4645

Qy   1679 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGGCA 1738
          ||||||||||||||||||||||||||||||||||||||||||||||  |||||||||||||
Db   4646 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCGCC 4705

Qy   1739 AGATGCCGAGGCCATGAGCCAGGAG---CCACCAGCCGCCAACATCCTGATCAGCCGCAGCAACT 1798
          |||||||||||||||||||||||||   ||||||||||||||| |||||| ||| |||||||||
Db   4706 GCATGAAGGGCCCCACCACCACACCCCCCACCAGCCGCGTGACCAAACATGCAGCGCGGCGCAACT 4765

Qy   1799 CCATGGAGAGCCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCATCCAGAAGG 1858
          |||  |||||||||||||| ||||||||||||||||||||||||||| ||||||||||||
Db   4766 CCACCAACCAGCGCATCGTGATCTGGGGCAAGGTTCAAGCTGCCATCCAGAGAGG 4825

Qy   1859 AGACCTGGGAGGACCTGGTGGACCGCCACTGGCACGCCGATCCCCGAGTGGGGAGT 1918
          ||||||||||||||||||||||||| |||||||| |||||||||||||||||||||
Db   4826 AGACCTGGGAGGCCTGGTGGACCGAGTACTGGCAGGATCCCTGATCCCCGAGTGGGAGGT 4885

Qy   1919 TCGTGAACACCCCCCCCCCCGTGTGAACGGCGCCCCAACCAGCTGGAAGAGAGCCCATCATCG 1978
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4886 TCGTGAACACCCCCCCCCCCGTGTGAACGGCGCCCCAGCTGGAAGAGAGCCCATCATCG 4945

Qy   1979 GCGCCGAGACTTCACGTGGACGGCGCCCGCAACCGCGAGACCAAGATCGGCAAGGCCG 2038
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4946 GCGCCGAGACTTCACGTGGACGGCGCCGCCAACCGCGAGACCAAGTGGGCAGGCCG 5005

Qy   2039 GCTACGTGACCGACCGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGA 2098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5006 GCTACGTGACCGACCGGGCCGGCCGCCAGAAGGTGTGCCCCTGACCGAGACCACCAACCAGA 5065

Qy   2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCG 2158
          ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   5066 AGACCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACAGCAGCGGCCTGAAGGTGAACATCG 5125

Qy   2159 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCTGACAAGAGGAGCGG 2218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5126 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCCAGAGAGCG 5185

Qy   2219 AGCTGGTGAACCAGATCATCGGCCGGCAACGAGCAGATGACAAGCTGTACCTGAGCTGGG 2278
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5186 AGCTGGTGGAGCCAGATCATCGAGCAGTGCGGCCAGCAGGTGTGTACCTGCCTGGG 5245

Qy   2279 TGCCCGCCCAAGGACATCGGCCGGCAACGAGAGCAGATCGACAAGCTCGGTGAGCAAGGGCA 2338
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5246 TGCCCGCCCAAGGACATCGGCCAAGCTGGTGCCGGCCATCGAGAAGGACCGGAA 5305

Qy   2339 TCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGAGGAGCACGAGAA 2393
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5306 TCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGAGGAGCACGAGAA 5360
```

RESULT 10
AX427927

```
LOCUS       AX427927                12411 bp    DNA        linear      PAT 20-JUN-2002
DEFINITION  Sequence 165 from Patent WO0232943.
ACCESSION   AX427927
VERSION     AX427927.1  GI:21538014
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE     Modifications of hiv env, gag, and pol enhance immunogenicity for
            genetic immunization
  JOURNAL   Patent: WO 0232943-A 165 25-APR-2002;
            GOVERNMENT OF THE UNITED STATES (US)
FEATURES             Location/Qualifiers
     source          1..12411
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match             82.1%;  Score 2027;  DB 6;  Length 12411;
Best Local Similarity   91.6%;  Pred. No. 6.3e-204;
Matches 2194;  Conservative    0;  Mismatches 185;  Indels   16;  Gaps    4;

Qy     14 TGGCCGAGGCCATGAGCCAGG---CCACCAGCCGCCAACATCCTGATCAGCCGCAGCAACT 70
          ||||||||||||||||||||||   ||||||||||||| |||||| ||| |||||||||||
Db   2967 TGGCCGAGGCCATGAGCCAGGCCAGCGTGACCAGCCGCCACCATCATGCAGCGCGGCGCAACT 3026

Qy     71 TCAAGGGCCCCACCAAGGCCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCC 130
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3027 TCCAACCAGCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACAACCGCCCC 3086

Qy    131 GCAACTGCGGCGCCCCCCCCGAAGAGGGGGCTCGGAGTCGGCGGAGGAGGGGCCACCAGA 190
```

```
Qy   2339 TCCGCAAGGTGCTGTTCCTGACGCGCATCGATGGCGGCATCGTGATCTACCAGTA 2393
          ||||||||||||||||||||||||||||||||||||||||||||   |||||||
Db   5306 TCCGCAAGGTGCTGTTCCTGACGCGCATCGATGGCGGCATCGTGATCTACCAGTA 5360
```

RESULT 11
AX427926
LOCUS   AX427926   9194 bp   DNA   linear   PAT 20-JUN-2002
DEFINITION   Sequence 164 from Patent WO0232943.
ACCESSION   AX427926
VERSION   AX427926.1 GI:21538013
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE   Modifications of hiv env, gag, and pol enhance immunogenicity for
             genetic immunization
  JOURNAL   Patent: WO 0232943-A 164 25-APR-2002;
             GOVERNMENT OF THE UNITED STATES (US)
FEATURES   Location/Qualifiers
     source   1..9194
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match   82.0%; Score 2025.4; DB 6; Length 9194;
Best Local Similarity 91.6%; Pred. No. 1e-203;
Matches 2193; Conservative   0; Mismatches 186; Indels   16; Gaps   4;
```

```
Qy     14 TGGCCGAGGCCATGAGCCAGG---CCACCAGCGCCAACATCCTGATGAGCGCAGCAACT   70
          |||||||||||||||||||||   ||||||||||||||   |||||||||||||||||
Db   2967 TGGCCGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATCAGCGCAGCCCACATCGCACT 3026
```

```
Qy     71 TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGC                      130
          |||||||||||||||||||||||||||||||
Db   3027 TCCCGCAACAGCGCCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAAGGGCCACACCGCC 3086
```

```
Qy    131 GCAACTGCCGCGCGCCCCCCGCCAAGCAGGG                             190
          |||||||||||||||||||||||||||
Db   3087 GCAACTGCCGCGCGCCCCCCGCCAAGCAGGGCGGCGGCCAAGGAGGCGGCCACCAGA     3146
```

```
Qy    191 TGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGG   250
          ||||||||||||||||||||||||||||||       |||||||||||||||
Db   3147 TGAAGGACTGCACCGAGCGCCAGCCGGCTAA-TTTTTTAGGGAAGAATCTGGCCTTCCACAAG 3205
```

```
Qy    251 GCAAGGCCCCCGCAGTTCCCCCGCAGCAGAACCGGCGCCAACAGCCCCAACAGCCCCGAGC 310
          |||||||||||||||||||||||||||||||||
Db   3206 GGAAGGCCAGGGAATTTTCTTCAGAGCAGACCGGCCGCCAACAGCCCCAACAGCCCCGAGAGC 3265
```

```
Qy    311 TGCAGTGCGGG------CGACAACCCCGCAGGCGGCGGCGGAGCGGCCGAGCGCCAGGGCA  364
          ||||||||||      ||||||||||||||||||||
Db   3266 TTCAGGTTTGGGAAGAGACAACAACTCCTCTCAGAAGCAGCAGGACCGATAGAACAAGGAA 3325
```

```
Qy    365 ------CCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCGTGTGAGCATCAAGG   418
                |||||||||||||||||||||||||||||||||||
Db   3326 CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGCAGCGACCGCCGGCAACACCGTGCTGCTGGAGG 3385
```

```
Qy    419 TGGGGGCCAGATCAAGGAGCCCGTCGGAAGCGACACCGGCGCCGCCAATGATCGGCGGCTTCATCA 478
          ||||||||||||||||||||||||
Db   3386 TAGGGGCCAGCTGACGATCAAGGAGCCCGTCGGGAGCGGCATCGGCGGCATCGGGGGCTTCATCA 3445
```

```
Qy    479 AGATGAGCGTCCCCGCAGATCAAGGGCCAGATCAAGGAGCC                    538
          |||||||||||||||||||||||||||||||||||
Db   3446 AGATGAACCTGGACCTGCCAGTGCCCAGATCAAGGAGCC                      3505
```

```
Qy    539 AGGTGCGCCCAGTACGACCCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   3506 AGGTGGGCCCAGTACGACCCAGATCCTGATCGAGATCTGCGGCAAGAGCCCAGGGCCATCGGCACCG 3565
```

```
Qy    599 TGCTGATCGGCCCACCCCGTGACATCATCGGCGGCAACATCGTCGACATGTCGACCCAGCTGG 658
          |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||
Db   3566 TGCTGGTGGGCCCACCCCGTGACATCATCGGCGGCAACATCCTGTCGACCCAGATCGGCT 3625
```

```
Qy    659 GCACCCTGAACTTCCCCATCAGCCCCATCAGCCCCATCGAGACCGTGCCCGTGACCGTGC    718
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3626 GCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGACCGTGGCGGCCGGCA 3685
```

```
Qy    719 TGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCG    778
          |||||||||||||||||||||||||||||||||||||||
Db   3686 TGGACCGGCCCAAGGTGAAGCTGTGCCCCCTGACCGAGGAAGATCAAGCCCCTGGTGG 3745
```

```
Qy    779 CCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCT 838
          |||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3746 AGATCTGCACCGAGGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCT 3805
```

```
Qy    839 ACAACACCCCGTGTTCGCCATCAAGAAGAAGACAGCAGCAGCCAAGTGGCGACACCTGGTGG 898
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3806 ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCAGCCAAGTGGCGCAAGTGGTGG 3865
```

```
Qy    899 ACTTCCGCGAGCTGAACAAGGCACCCGGGACTTCGTGGAGGTGCAGCTGGGCATCCCCC 958
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3866 ACTTCCGCGAGCTGAACAAGGCACCCGGGACTTCGTGGAGGTGCAGCTGGGCATCCCCC 3925
```

```
Qy    959 ACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCCTACT 1018
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3926 ACCCCGCCGGCCTGAAGCAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCCTACT 3985
```

```
Qy   1019 TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCCGCCTTCACCATCCCCAGCATCA 1078
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3986 TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCCGCCTTCACCATCCCCAGCATCA 4045
```

```
Qy   1079 ACAACGAGACCCCGGCATCCCGGCATCCGCTACCAGTACAACGTGCTGCCCAGGGTGGGAGGGCA 1138
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4046 ACAACGAGACCCCGGCATCCCGGCATCCGCTACCAGTACAACGTGCTGCCCAGGGCTGGAAGGGCA 4105
```

```
Qy   1139 GCCCCAGCATTTCCAGAGCAGCAGATCCTGGAGCCCTTCCGCGCCCAACC          1198
          |||||||||||||||||||||||||||||||||||||||||
Db   4106 GCCCCGCCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGAAGCAGAAACC 4165
```

```
Qy   1199 CCGAGATCGTGATCTACCAGTACATGACATGGAGGAGCGCGACCGTGACTGTGGCGAGATCG 1258
          |||||||||||||||||||||||||||||||||||||||||||||
Db   4166 CCGACATCGTGATCTACCAGTACATGACATGGAGGAGCGCGACCACCTGTACGTGGGCAGCGGATCG 4225
```

```
Qy   1259 GCCAGCACCGCGCGCCAAGATCCAAGATCGAGGAGCTGCGCCAAGCAGACCTGCTGCGCGGCTTCACCA 1318
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4226 GCCAGCACCGCGCGCCAAGATCCAAGATCGAGGAGCTGCGCCAAGCAGACCTGCTGCGCGGCTTCACCA 4285
```

```
Qy   1319 CCCCCGACAAGAGCACCAGAGAGGAGCCCCTTCCTGTGATGGGCTACGAGCTGCACC 1378
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4286 CCCCCGACAAGAGCACCACCAGAGAGGAGCCCCCTTCCTGTGATGGGCTACGAGCTGCACC 4345
```

```
Qy   1379 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACG 1438
          ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db   4346 CCGACAAGTGGACCGTGCAGCCCGTCGTGCCCGAGAAGGAAGGAGCAGTGGACCGTGAACG 4405
```

```
Qy   1439 ACATCCAGAAGTGGTGGGCAAGCTGGGCCAGCCAGCCCAGATCTACCCCGGCCATCAAGG 1498
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4406 ACATCCAGAAGCTGGTGGGCAAGCTGGGCCAGCCAGCCGACTACCGCGGCATCAAGG 4465
```

```
Qy   1499 TGCCCAGCTGTGCAAGCTGCTGCGAGCTGGCCGAGAAGGCCCCCAAGGCCCCGACATCGTGCCCCTGA 1558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4466 TGCGCCAGCTGTGCAAGCTGCTGCGAGCTGGCCGAGAAGGCCCCAAGGCCCTGACGAGGTGCCCCTGA 4525
```

```
Qy   1559 CCGAGGAGGCCGAGATCAAGGAGCCCCTTCTGACACACACCGGCGCCGCAGCCGTGCACG 1618
          |||||||||||||||||||||||||||||||||||||||||||
Db   4526 CCGAGGAGGCCGAGATCAAGGAGATCCTGGAGATCCTGCTGAAGGAGGCCCGTGCACG 4585
```

```
Qy   1619 GCGTGTACTACGACCCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCACC 1678
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4586 GCGTGTACTACGACCCAGATCCTGATCGAGATCTGCGGCAAGAGCCCAGGGCCATCGGCACCG 4645
```

```
KARIN HARTOG,
  PI  HONG LIU,CATHERINE GREER,MARK SELBY,CHRISTOPHER WALKER PC
C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P31/18,A61P37/02, PC
C12N5/10,
  PC  C12N7/00,C12P21/02,C12N15/00,C12N5/00,A61K37/02 CC
Description of Artificial Sequence: FS(-).protmod.RTopt.YM FH  Key
                        Location/Qualifiers                   .protmod.RTopt.YM FH
                        1..2306
FEATURES        source
  FT  source         Location/Qualifiers
  FT                 1..2306
                     /organism='Artificial Sequence'.
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"

ORIGIN

Query Match    82.0%;  Score 2025.2;  DB 6;  Length 2306;
Best Local Similarity  93.3%;  Pred. No. 1.5e-203;
Matches 2155;  Conservative  0;  Mismatches 138;  Indels  18;  Gaps  3;
```

```
Qy   170 GCGGCCAAGAGAGGGCCACCAGATGAAGGACTGACCGAGCGCGCCAAGGCCAACTTCTTCCGCG 229
             ||||||||||||  |||||||  ||||||||||||||||||||| ||||||||||  || |||
Db     1 GCGGGCGCGAAGGACACCCAAATGAAGATTGCACTGAGAGACAGGCTAATTTCTTCCGCG  60

Qy   230 AGGACCTGGCCTTCCCCCAGCCAGGGCCCGCCGAGTTCCCCAGGGCAGGACCGAGAACCGCGCCA 289
             |||||||||||||||||||  |||||||||||||||||| |||||||||||||||||||||||
Db    61 AGGACCTGGCCTTCCTCTGCAGGGCAAGGCCCGCAGTTCAGCAGCAGGCCCGCAGCCGCGCCA 120

Qy   290 ACAGCCCCACCAGCCGCGCCGAGCTGCGAGGTGCGCGGCG------ACAACCCCGCCAGCGAGG 343
             |||||||||||||||||||||||||||||                    |||||||||||||
Db   121 ACAGCCCCACCAGCCGCGCCGAGCTGCGAGGTGTGGGCGGCGAGAACAACAGCCTGAGCGAGG 180

Qy   344 CCGGCGCCAGCGCCAGCGCCACCCTG------AACTTCCCCAGATCACCCGTGTGGCAGC 397
             |||||||||||||||||||||||||         ||||||||||||||||||||||||||
Db   181 CCGGCGCCAGCGCCAGCGCACCGTGAGCTTCAACTTCCCCAGATCACCCGTGTGGCAGC 240

Qy   398 GCCCCCTGGTGAGCATCAAGGTGGCGGGCCTGCTGCAGATCAAGGAGCCCTGCTGAGACCGGCG 457
             |||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db   241 GCCCCCTGGTGACCATCAGCAGCGGCATCGGCGGAGATCAAGGAGCTCAAGGAGCCCTGCTGAGACCGGCG 300

Qy   458 CCGACACACCGTCTGCTGGAGGATGAGCTGCCCGGCAAGCCAAGATCGAGACCTGAGATGATCG 517
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 CCGACACACCGTGGAGGAGGATGAAACTGCCCGGCAAGCCAAGACCAGATCCGAGATGATCG 360

Qy   518 GCAAGGATCGGCGGCTTCATCAAGGTGCGCGGCGATCGGCCCTGAATCATCGGCCGCCAGA 577
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GCGGGATCGGCGGCTTCATCAAGGTGCGCGGCGATCGGCCCTGAACAATCATCGGCCGCCAGA 420

Qy   578 GCAAGAGGCCATCGGCACCGGCATCGGCACCGGCTGCTGCTGGTGGCCCACCCCGTGAACATCATCGGCCGCCA 637
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCCACAAGGCCATCGGCACCGGCACCGGCTGCTGCTGGTGGCCCACCCCGTGAACATCATCGGCCGCCA 480

Qy   638 ACATGCTGACCCAGCTGGCTGCACCCTGAACTTCCCCATCAGCCCATCGGAGACCGTGC 697
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ACCTGCTGACCCAGCTGGCTGCACCCTGAACTTCCCCATCAGCCCCATCAGCGGAGACGGTGC 540

Qy   698 CCGTGAAGCTGACGGCCCGGCATGACGGCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG 757
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCGTGAAGCTGACGGCCCGGCATGACGGCCCAAGGTGAAGCAGTGTGGCCCCTGACCGAGG 600

Qy   758 AGAAGATCAAGGCCCTGACCGCCATCTGCGACGAGATGAGGAGATGAGAAGGCAAGATCACCA 817
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGAAGATCAAGGCCCTGGTGGAGATCGTCACCGAGATGGAGGAAGGAGGGCCAAGATCAGCA 660

Qy   818 AGATCGGCCCCGAGAACCCCTACACACCCCGTGTTGCCATCAAGAGAAGAGACAGCA 877
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AGATCGGCCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGAACAGCA 720

Qy   878 CCAAGTGGCCAAGCTGGTGGCACTTCGCGACTGAACAAGGCCACCCAGGACTTCTGGG 937
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCAAGTGGCCAAGCTGGTGGCACTTCGCGGACTGAACAAGGCCGCACCCCAGGACTTCTGGG 780

Qy   938 AGGTGCAGCTGGGCATCCCCCACCCCGCCGCGCGCGCCGGGAAGAGAAGAGAAGAGAGCGTGACCGTGC 997
```

```
Qy  1679 AGTGGACCTACCAGGAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCCAAGTACGGCCA 1738
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4646 AGTGGGACCTACCAGGAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCCAAGTACGGCCC 4705

Qy  1739 AGATGGCACCGCACCCACCAACACGACGTGAGCGACCAGCTGACCTGAGCGGCCGTGCAGAGATCG 1798
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4706 GCATGAAGGGCGGCGCCCACCAACCCAACGACGTGGAGCGACCAGCTGGAGCAGCAGGCCGTGCAGAGATCG 4765

Qy  1799 CCATGGAGAGCATCGTGTGATCTGGGCAAGACCCCCAGTTCCGCCTGCCATCCAGAAGG 1858
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4766 CCACCGAGAGCATCGTCGTGATCTGGGCAAGGCAAGCCCCCAAGTTCAAGCTGCCCATCCAGAAGG 4825

Qy  1859 AGACCTGGGAGACCTGGTGACCGACTACTGCAGGCCACCTGGATCCCCGAGTGGGAGT 1918
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4826 AGACCTGGGAGGACCTGTGGAGCCTGGTGGACCGACCGACTGCAGGCCACCTGGATCCCCGAGTGGGAGT 4885

Qy  1919 TCGTGAACACCCCCCCCTGTGAAGTGTGTACCAGCTGGTGAGGAGGAGCGCCATCATCG 1978
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4886 TCGTGAACACCCCCCCCCCCTGGTGAAGCTGTGTACCAGCTGGTGAGGAGGAGCGCCCATCATCG 4945

Qy  1979 GCGCCAGAGACCTTCACGTGGACGGGCCGCCAACCGCGAGACCAAGATCGGCAAGGGCG 2038
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4946 GCGCCGAGACCTTCTACGTGGACGGGCGCGGCCAACCGCGAGACCAAGGTCGGCAAGGGCCG 5005

Qy  2039 GCTACAGTGACCGACCGGGGCCCGGCAGAAGATCGTGAGCCTGACCGACCACCACCAGA 2098
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5006 GCTACAGTGACCGACCGGGGCCCGGCAGAAGATCGTGAGCTCGACCAGACCACCACCAGA 5065

Qy  2099 AGACCGAGCTGCGAGCCATCCAGCTGGCCCTCGCAGGACAGCCCGCAGCGTGAACATCG 2158
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5066 AGACCGAGCTGCGAGCCATCCAGCTGGCCCTCGCAGGACAGCCGCAGCGGTGAACATCG 5125

Qy  2159 TGACCGACAGCCAGTACGCCCTTGGGCATCACCAGGCCCGACCAGAGCGGAGAGCG 2218
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5126 TGACCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCCGACACAAGAGCGGAGAGCG 5185

Qy  2219 AGTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG 2278
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5186 AGTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCCTGGG 5245

Qy  2279 TGCCCGCCCAAGGGCATGGCAAGATCGGCGGGAACAGCGAGATCGACAAGGTGGTCAAGGGCA 2338
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5246 TGCCCGCCCCAAGGGCATGGCATCGGCGGGAACAGCGAGATCGACAAGGTGGCGCGGGCA 5305

Qy  2339 TCCGGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGCCCAGGAGGACACGAGAA 2393
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5306 TCCGGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGGCCCAGGAGGACACGAGAA 5360
```

```
RESULT 12
BD263704
LOCUS        BD263704               2306 bp    DNA        linear      PAT 17-JUL-2003
DEFINITION   Improved expression of HIV polypeptides and production of
             virus-like particles.
ACCESSION    BD263704
VERSION      BD263704.1  GI:33073472
KEYWORDS     JP 2002533124-A/71.
SOURCE       synthetic construct
  ORGANISM   synthetic construct
             artificial sequences.
REFERENCE    1  (bases 1 to 2306)
  AUTHORS    Barnett,S., Megede,J.Z., Srivastava,I., Lian,Y., Hartog,K., Liu,H.,
             Greer,C., Selby,M. and Walker,C.
  TITLE      Improved expression of HIV polypeptides and production of
             virus-like particles
  JOURNAL    Patent: JP 2002533124-A 71 08-OCT-2002;
             CHIRON CORP
COMMENT      OS   Artificial Sequence
             PN   JP 2002533124-A/71
             PD   08-OCT-2002
             PF   30-DEC-1999 JP 2000591193
             PR   31-DEC-1998 US  60/114495,01-DEC-1999 US  60/168471 PI
             SUSAN BARNETT,JAN ZUR MEGEDE,INDRESH SRIVASTAVA,YING LIAN, PI
```

```
Db  781 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        AGTGCAGCTGGGCATCCCCACCCCGCCGGCCTGAAGAAGAAGAGGAGGCGTGACCGTGC  840

Qy  998 TGGACGTGGGCGACGCGCTACTTCAGGTGCCGCCTGCCCCTGGACGACTTCCGCAGTACACCG  1057
        |||||||||||||||||||||||||||||| |  |||||||||||||||
Db  841 TGGACGTGGGCGACGCGCTACTTCAGGTGCCGCCTGCCCCTGGACAACGACTTCCGCAGTACACCG  900

Qy 1058 CCTTCACCATCCCAGACATCAACAACAGAGACACCCCGGCATCCGCTACCAGTACAACGTGC  1117
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 CCTTCACCATCCCAGACATCAACAACAGAGACACCCCGGCATCCGCTACCAGTACAACGTGC  960

Qy 1118 TGCCCCAGGGCTGTGAAGGGCAGCCCAGCAGCATCTTCAGAGACAGCATGACCAAGATCCTGG  1177
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db  961 TGCCCCAGGGCTGGAAGGGCAGCCCGCCATCTTCAGAGACAGCATGACCAAGATCCTGG  1020

Qy 1178 AGCCCTTCCGCGCCGCGACCCGAACCCGACATCGTGATCTACCAGTACCAGTGATCACATGGACGACCTGTACG  1237
        ||||||||||||||||||||||||||||  |||||||||||-------||
Db 1021 AGCCCTTCCGCAAGCAAGCAGAGACCCGACATCGTCACCA------GGCCCCCCCTGTACG  1074

Qy 1238 TGGCAGCGACCTGGAGATCGGCCAGACCCGCGCCAGATCGAGCGAGTCTGAGCTGCGCAAGCACC  1297
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1075 TGGGCAGCGACCTGGAGATCGGCCAGCACCACCCGCCAAGATCGAGCGAGTCTGCGGCAGACC  1134

Qy 1298 TGTGCGGCTGGGCTTCACCACCCCGACAAGGAACACCAGAAGAGCACCCCCCTTCCTGT  1357
        ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1135 TGCTGCGGCTGGGCTTCACCACCCCGACAAGAAGCAGAGACACCCCCCCCTTCCTGT  1194

Qy 1358 GGATGGGCTACGAGCTGCACGACCTGCACCGTGCCCATCGAGCTGCCCGAGA  1417
        ||||||||||||||||||||||||||||||||||||||||||
Db 1195 GGATGGGCTACGAGCTGCACGACCTGCACCGTGCAGCCATCATGCTGCCCGAGA  1254

Qy 1418 AGGAGAGCTGGACCGTGAACGACATCCAGAACGTGACTGGGTGGGCAAGCTGAACTGGCCAGCC  1477
        |||||||||||||||||||||||||||||||||||||
Db 1255 AGGACAGCTGGACCGTGAACGACATCCAGAACTGGTGCGGCAAGCTGAACTGGGCCAGCC  1314

Qy 1478 AGATCTACCCCGGCCATCAAGGTGCGCCAGCTGCTGCGCGGGCAAGCCC  1537
        ||||||||||||||||||||||||||||||||||||||||
Db 1315 AGATCTACGCCGGCCATCAAGGTGAAGCGCAGCTGTGCGGCAGCCCC  1374

Qy 1538 TGACCGACATCGTGCCCCGAGCTGCCGAGCTGGCCGAGACCGGCGAGA  1597
        ||||||||||||||||||||||||||||||||||
Db 1375 TGACCGAGGTGATCCCCCGACGAGGCTGCGAGCTGGGCCGAGACCGGCGAGA  1434

Qy 1598 TCCTGCGGGAAGGCCGTGCCACGGCGTGTACTACGGACCCCCAGCAAGAAGGACCTGGTGGCCGAGA  1657
        |||||||||||||||||||||||||||||||||||||||||||||||
Db 1435 TCCTGAAGGAGGCCCGTGCCACGAGGTGTACTACGGACCCCCAGCAAGAAGGACCTGGTGGCCGAGA  1494

Qy 1658 TCCAGAAGCAGGGGCCACGACCAGTGACCTACCAGATCTACCAGAGGAGCCCTTCAAGAACC  1717
        ||||||||||||||||||||||||||||||||||
Db 1495 TCCAGAAGCAGGGGCCCAGGTGACCTACCAGATCTACCAGAGGAGCCCTTCAAGAACC  1554

Qy 1718 TGAAGACCGGCAAGTACGGCCAAGATGCGCACCGCCACCACCGCGCCACCAACGACGTGAAGCAGCTGA  1777
        |||||||||||||||||||||||||||||||||||||||
Db 1555 TGAAGACCGGCAAGTACGGCAAGATCGGCGCCATGCCGGCCGCCGCGCCACCAACGACGTGAAGCAGCTGA  1614

Qy 1778 CCGGGCAAGGAGGGCCCAGCACCCAGGCAGAGATGAAGATCGGCCAACTTCTTCCGGG  1837
        |||||||||||||||||||||||||||||||
Db 1615 CCGAGGGCCCGAGGGACACCACCAGGCCAAGATCCGTGATCGGCCAAGATCCCCAAGT  1674

Qy 1838 TCCGCCTGCCCATCGAAGGAGAGGACCTGGAGACCGACTACTGCCAGGCCA  1897
        ||||||||||||||||||||||||||||||||
Db 1675 TCAAGCTGCCCATCCAGGAGGAGCCCTGGGATGGAGTGGTACTGCGCCAGGCCA  1734

Qy 1898 CCTGGATCCCCGGAGTTGGGAGTTCGTGAACACAACCCCCCAGCTGGTGAAGCTGTGGTACCAGC  1957
        |||||||||||||||||||||||||||||||||||||||||||
Db 1735 CCTGGATCCCCGAGTTGGGAGTTCGTGACAACACCCCCCAGCTGGTGAAGCTGTGGTACCAGC  1794

Qy 1958 TGGGAGAAGAAGGAGAGCCCCATCATCGGCGCCATCATCGTGTGGGGCCAGCCCGCCGCCAACCGCG  2017
        ||||||||||||||||||||||||||||||||||||||
Db 1795 TGGAGAAGAAGGAGAGCCCATCATCGTGTGGGGCGCCGAGATCGTGAGCGC  1854

Qy 2018 AGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGCCCCGACCGGCGCAGAGATCGTGAGCGC  2077
        ||||||||||| ||||||
```

```
Db 1855 AGACCAAGCTGGGCAAGGCCGGCTACGTGACCGACCGGGCCGCCCGGCCGAGAGGTGTGAGCA  1914

Qy 2078 TGACCGAGACCAACCACCAGAGAAGACCGAGCTGCAGGCCATCCAGCTGCGGCCTGCGCTCCAGGGACA  2137
        ||||||||||||||||||||||||||||||||||||||||
Db 1915 TCGCCCACACCACCAACCACCAGAGAAGACCGAGCTGCAGGGCCATCCACCTGCGCCTGCGCCGAGGACA  1974

Qy 2138 GCGGCAGCGGAGGTGAACATGTGACCGGACGCAGCAGCAGTACGGCCCTGGCATCATCCAGGCCC  2197
        |||||||||||||||||||||||||||||||||||||||
Db 1975 GCGGCCTGGAGGTGAACATCGTGACCGGACGCAGTACGGCCCTGGCATCATCCAGGCCC  2034

Qy 2198 AGCCCGACAAGAGAGGAGAGCGAGCTGGTGAACCAGATCATCGGCGCAGCTGATCAAGAAGG  2257
        ||||||||||||||||||||||||||||||||||||||||||||
Db 2035 AGCCCGACAAGAGAGGAGAGCGAGCTGGTGAGCCAGATCATCGGCGCAGCTGATCAAGAAGG  2094

Qy 2258 AGAAGTGTACCTGAGCTGGGGTGGGTGCCCCGGCCCACAAGGGCATCGGGGCAACGAGCAGATCG  2317
        ||||||||||||||||||||||||||||||||||||||||||||
Db 2095 AGAAGTGTACCTGCCTGGGGTGCCCCGGCCCACAAGGGCATCGGGGCAACGAGCAGGTGG  2154

Qy 2318 ACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACCGGCATCGATGGCGGCA  2377
        |||||||||||||||||||||||||||||||||||||||||||
Db 2155 ACAAGCTGGTGAGCGCCGGCATCCGCAAGGTGCTGTTCCTGAACCGGCATCGATGGCGGCA  2214

Qy 2378 TCGTGATCTACCAGTACATGACGACCTGTACGTGGGCAGCGGCCGCGCCCTAGGATCGATT  2437
        |||||||||||||||||||||||||||||||||||||||||
Db 2215 TCGTGATCTACCAGTACATGACGACCTGTACGTGGGCAGCGGCCGCGCCCTAGGATCGATT  2274

Qy 2438 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC  2469
        |||||||||||||||||||||||||||
Db 2275 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC  2306
```

RESULT 13
```
AR373387
LOCUS       AR373387         2306 bp    DNA        linear    PAT 18-DEC-2003
DEFINITION  Sequence 82 from patent US 6602705.
ACCESSION   AR373387
VERSION     AR373387.1  GI:40075490
KEYWORDS    Unknown.
SOURCE      Unknown.
  ORGANISM  Unclassified.
REFERENCE   1  (bases 1 to 2306)
  AUTHORS   Barnett,S.W., Megede,J., Greer,C. and Selby,M.
  TITLE     Expression of HIV polypeptides and production of virus-like
            particles
  JOURNAL   Patent: US 6602705-A 82 05-AUG-2003;
FEATURES             Location/Qualifiers
     source          1..2306
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match         82.0%;  Score 2025.2;  DB 6;  Length 2306;
Best Local Similarity   93.3%;  Pred. No. 1.5e-203;
Matches 2156;  Conservative  0;  Mismatches 138;  Indels 18;  Gaps 3;
```

```
Qy  170 GCGGGCAAGGAGGGCCACCAGATGAAGGACTCCACCCAGCGGCCAACTTCTTCCGGG  229
        |||||||||||||||||||||||||||  ||||||||||||||||||||
Db    1 GCGGGCCCGAGGGACACCACCAGGCCAAATGAAAGATTGCACTGAGAGAAGATTCTTCCGGG  60

Qy  230 AGGACTGGCCTTCCCCAGGGCCAAGGTTCCCCGAGTTCCCCGAGTTCAGCGAGCAGAACCGCGCCA  289
        ||||||||||||||||||||||||||||||||||||||||||
Db   61 AGGACCTGGCCTTCCTCCTGCAGGGCCAAGGGCCCCGAGTTCAGCGAGCAGCAGACCCGCGCCA  120

Qy  290 ACAGCCCGAGCCGCAGCCGCGCGGG------ACAACCCCCGAGCAGGCGAGG  343
        |||||||||||||||||||||||||||||||  ||||||||||||
Db  121 ACAGCCCCACCCCGAGCCGCAGTGTGCCGAGAACACGACCTGAGCGGGAGG  180

Qy  344 CCGGGCGCGAGCCGCAGGGCACCTG-----AACTTCCCCCAGATCACCCTGTGGCAGC  397
        |||||||||||||||||||||||||     ||||||||||||||||||||
Db  181 CCGGGCGCGACCCGACCCCAGATCTTCAACTTCCCCCAGATCACCCTGTGGCAGC  240

Qy  398 GCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGG  457
        |||||||||||||||||||||||||||||||||||||||||| |||||||
```

Left column (coordinate pairs shown at ends of rows):

```
Db   241  ...GCCCCCTGGTGACCATCAGGATCGGGGCGCCAGCTCAAGGAGGAGGCGCTGCTCGACACCGGCG  300
Qy   458  CCGAGGACACCGTGCTGGAGAGATGAGCGTGCCCGGCGAATGGAAGCTGAAGCCCAAGATGATCG  517
Db   301  CCGACGACACCGTGCTGGAGAGATGAACCTGCCCGGCGAATGGAAGCTGAAGCCCAAGATGATCG  360
Qy   518  GCGGGCATCGGGGGCTTCATCAAGGTGCGCGAGTAGGACCAGATCCTGATCGAGATCTGGCG  577
Db   361  GCGGGCATCGGGGGCTTCATCAAGGTGCGCGAGTAGGACCAGATCCCGTGGAGATCTGCCG  420
Db   421  GCCACAAGGCCATCGGGCACCGTGCTGGTGCTGGTGGGCCCCACCCCGTGAACATCATCGGCCGCA  480
Qy   578  GCAAGAGGCCATCGGCACCGTGCTGTGATCGGCGCCCACCCCGTGAACATCATCGGCCGCA  637
Qy   638  ACATGCTGACCCAGCTGGGGTTGCACCTGACACTTCCCATCAGCCCCATCGAGACCGTGC  697
Db   481  ACCTGCTGACCCAGATCGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACGGTGC  540
Qy   698  CCGTGAAGCTGAAGCCCGGCATGGACGGGCCCCAAGGTGGCGCCCATCTGGGCCCTGACCGAGG  757
Db   541  CCGTGAAGCTGAAGCCCGGGATGGACGGGCCCCAAGGTCAGCTCAGCGGCCCTGACCGAGG  600
Qy   758  AGAAGATCAAGGCCCTGACCGCCCATCTGCGAGGAGATGAGAAGGAGGGCAAGATCACCA  817
Db   601  AGAAGATCAAGGCCCCTGGTGGAGATCTGCACCGAGAGGTGGGAGAAGGGCAAGATCAGCA  660
Qy   818  AGATCGGCCCCGAGAACCCCTACAACACACCCCGTGTCGCCATCAAGAAGAAGGACAGCA  877
Db   661  AGATCGGCCCCGAGAACCCCTACAACACACCCGTGTCGCCATCAAGAAGAAGGACAGCA  720
Qy   878  CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGG  937
Db   721  CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGG  780
Qy   938  AGGTGCAGCTGGGCATCCCCCACCCCCAGCTTCAGGCTGCCCTGAAGAGAAGAGCGTGACCGTGC  997
Db   781  AGGTGCAGCTGGGCATCCCCCACCCCCAGCTTCAGGCTGCCCTGAAGAGAAGAGCGTGACCGTGC  840
Qy   998  TGGACGTGGGCGCACGGCTACTTCAGGCTGCCCTGAGGAGACTTCCGCAAGTACACCG  1057
Db   841  TGGACGTGGGCGCACGGCTACTTCAGGCTGCCCTGAGGAGACTTCCGCAAGTACACCG  900
Qy  1058  CCTTCACCATCCCAGCATCAACAACAGAGCCCCGGCATCTTCGCTACCAGTACAAACGTGC  1117
Db   901  CCTTCACCATCCCAGCATCAACAACAGAGCCCCGGCATCCGCATCCTACCAGTACAAACGTGC  960
Qy  1118  TGCCCCAGGGCTGGAAGGGCGCTGGAAGGGCAGCATGACTCTTCCAGCAGCATGACCAAGAGATCCTGG  1177
Db   961  TGCCCCAGGGCTGGAAGGGCTGGAAGGGCAGCATGACCCAGCAGCATGACCAAGATCCTGG  1020
Qy  1178  AGCCCTTCCGCGCCCCGAGATCGTGATCCGAGATGCAGCGGCTCCAGTACCAGTCACGGGTTACG  1237
Db  1021  AGCCCTTCCGCGCCCCGAGATCGTGATCGTGATCTACCA------GGCCCCCCGTACG  1074
Qy  1238  TGGGCAGCGACCTGAGATCGGCCAGCAGATGCTGGCCAAGATCTGAGAGACGCTGCGCAAGCACC  1297
Db  1075  TGGGCAGCGACCTGAGATCGGCCAGCAGATCGAGGAGCTCGCGCCAGCACC  1134
Qy  1298  TGCTGCGGCTGGGGCTTCACCACCCGCCACCACCACCCCCTTCCTGT  1357
Db  1135  TGCTGCGGCTGGGGCTTCACCACCCGCCACCACCACCCCCTTCCTGT  1194
Qy  1358  GGATGGGCTAGAGCTGCACCCGACCAAGTGACCGTGCAGCTCGAGCTGCCCGAGA  1417
Db  1195  GGATGGGCTACCAGCTGCACCCGACCAAGTGACCCGTGCAGCTGCCCGAGA  1254
Qy  1418  AGGAGAGAGCTGACCGTGAACGACACATCCAGAAGAGCTGAACTGGGCCAGCC  1477
Db  1255  AGGACAGCTGACCGTGAACGACACATCCAGAAGAGCTGTGGGGCCAGCC  1314
Qy  1478  AGATCTACCCCGCCATCAAGGTGCGCCAGCTGTCGCCGGCCCAAGGCCC  1537
```

Right column (coordinate pairs shown at ends of rows):

```
Db  1315  AGATCTACGCCGGCATCAAGGTGAAGCAGCTGTGCGGCGGCCACCAAGGCCC  1374
Qy  1538  TGACCGACATCGTGCCCCTGACCGAGGAGCTGCGAGCTGGGCCGAGAACCGCGGAGA  1597
Db  1375  TGACCGAGGTGATCCCCCTGACCGAGGAGCTGCGAGCTGGGCCGAGAACCGCGGAGA  1434
Qy  1598  TCCTGCGCGGAGGCCGTGCACGGCCGTGTACTACGACCTGGTGCGCCGAGA  1657
Db  1435  TCCTGAAGAGGCCGTGCAGGAGGTGTACTACGACCCCAGCAGGACCTGGTGCGCCGAGA  1494
Qy  1658  TCCAGAAGCAGAGGCCACGACCAGTGGACCTACCAGGAGCCCTTCAAGAACC  1717
Db  1495  TCCAGAAGCAGGGCCAGGCCAGTGACCTACCAGATCTACCAGGAGCCCTTCAAGAACC  1554
Qy  1718  TGAAGACCCGCAAGTACGCCAAGATGCCAAGACGCCACCGCCCACCAACGACGTGAAGCAGCTGA  1777
Db  1555  TGAAGACCGGCAAGTACGCCCCATGCCCCACCAACGACGACCAGCAGCTGA  1614
Qy  1778  CCGAGGCCGTGCAGAAGCTGCATGGAAGATCGCCATGAGAGCATCGTGGGCAAGACCCCCAAGT  1837
Db  1615  CCGAGGCCGTGCAGAAGCTGAGCGAGAAGGAGACCTCGGTGATGGATGTGTACCAGC  1674
Qy  1838  TCCGCCTGCCATCCAGAAGAGACCTGGGACGCTGGTGGACCACTGCGAGGCA  1897
Db  1675  TCAAGCTGCCATCCAGGAGGAGACCTCGGGTGATGGATGTGTACCAGC  1734
Qy  1898  CCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCCTGGTGAACGCGGCCGCAACCGGCCG  1957
Db  1735  CCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCCTGGTGAACGCGGCCGCAACCGCG  1794
Qy  1958  TGGAGAAGGAGCCCATCATCGCGAGGCGGCTACGGCCGGCCGAGAGATCGTGAGCCG  2017
Db  1795  TGGAGAAGGAGCCCATCGTGGGCGCCGGCTACGGCCGCCGCAACCGCG  1854
Qy  2018  AGACCAAGATCGGCAGGAAGAAGAAGAGCCTGACCGGCGTACGGCCGGAGAAGATCGTGAGCGACA  2077
Db  1855  AGACCAAGCTGGGCGACCAGCGGCTACGGCCGGCCCGCAGAAGGTGGTGAGCA  1914
Qy  2078  TGACCGAGACCACCAACCAGAAGACCAGCTGCAGGCCGCATCCAGCTGCCCCTGCAGCACA  2137
Db  1915  TGCCGACACCACCAACCAGAAGACCAGCTGCAGGCCGCATCCAGCTGCCCCTGCAGCACA  1974
Qy  2138  GCGGCAGGCAGGTGAACATCGTACGCGACGCCAGTACGCCCCTGGGCCATCATCAGCGCC  2197
Db  1975  GCGGCCTGGAGGTGAACATCGTACGCGACGCCAGTACGCCCCTGGGCCATCATCAGCGCC  2034
Qy  2198  AGCCCGACAAGAGCGACGAGCGAGCTCATCGATCGAGCAGCTGATCAAGAAGG  2257
Db  2035  AGCCCGACAACAGAGCAGGAGCGAGCAGCTGGTGAGCCAGCAGCATCATCAAGAAGAGG  2094
Qy  2258  AGAAGGTGTACTTGAGCTGGTGGTCCCCCCACAAGGGCCGCCAACGACGCAGAGATCG  2317
Db  2095  AGAAGGTGTACCTGGCCTGGGTGGTCCCCCCACAAGGGCCGCATCCGGACGGCATCGATGGCGGCA  2154
Qy  2318  ACAAGCTGGTGAGCAGTACATGGACGACCTGTACGTGGACGGCCATCGATGGGCCGGCA  2377
Db  2155  ACAAGCTGGTGAGCAGTACATGGACGACCTGTTCCTGAACGGCGCATCGATGGGCCGGCA  2214
Qy  2378  TCGTGATCTACCAGTACACATGGACAGTCGGCAGCCGCCTAGGATCGATT  2437
Db  2215  TCGTGATCTACCAGTACACATGGACAGTCGGCAGCCGCCCTAGGATCGATT  2274
Qy  2438  AAAGCTTCCCGGGGCTAGCACCGGTGAATTC  2469
Db  2275  AAAGCTTCCCGGGGCTAGCACCGGTGAATTC  2306
```

```
RESULT 14
AX427938

LOCUS       AX427938      9785 bp    DNA    linear   PAT  20-JUN-2002
DEFINITION  Sequence 176 from Patent WO0232943.
ACCESSION   AX427938
VERSION     AX427938.1  GI:21538025
```

```
KEYWORDS   .
SOURCE     synthetic construct
ORGANISM   synthetic construct
           artificial sequences.
REFERENCE  1
AUTHORS    Huang,Y. and Nabel,G.J.
TITLE      Modifications of hiv env, gag, and pol enhance immunogenicity for
           genetic immunization
JOURNAL    Patent: WO 0232943-A 176 25-APR-2002;
           GOVERNMENT OF THE UNITED STATES (US)
FEATURES   Location/Qualifiers
  source   1..9785
           /organism="synthetic construct"
           /mol_type="unassigned DNA"
           /db_xref="taxon:32630"
           /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match        81.8%;   Score 2020.4;   DB 6;   Length 9785;
Best Local Similarity   91.6%;   Pred. No. 3.3e-203;
Matches 2190;   Conservative    0;   Mismatches 181;   Indels 21;   Gaps 4;
```

```
Qy      14  TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCGAGCTGCAGCGCGAGCAACTTCA   73
            |||||  |||||||||  ||||  |||||
Db    2957  TGGCCGAGGCCATGAGCCAGGCCATGAGCGCAGCACCACCATCATGATGCAGCGCGACAACTTCA   3026

Qy      74  AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCACCAGATGA   133
            ||||||  ||||
Db    3027  AGGGCCGCGCCCCGGCCCGG---CATCAAGTGCTTCAACTGCGGCAAGGAGGGGCACCAGATGA   3083

Qy     134  ACTGCCGCGCCCCGGCCCGGCCAGCGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCA   193
            |||||||||
Db    3084  ACTGCCGCGCCCCGGCCCGGCCAGCGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCA   3143

Qy     194  AGGACTGCACCGAGCGCCAGGCGACCAGGGCCAA------TAGGGAAGATCGGCCTTCCCCACAAGGGGA   253
            |||||||||||
Db    3144  AGGACTGCACCGAGCGCCAGGCCAGCCGGCCTAA------TAGGGAAGATCTGGCCTTCCCCACAAGGGGA   3197

Qy     254  AGGGCCGCGAGTTCCCCAGCGAGACCGCCCACCACGCGCGAGCTGC   313
            |||||||||
Db    3198  AGGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAGAGAGAGTTC   3257

Qy     314  AGTGTGCGCGG------CGACAACCCCGCAGCGGAGGCGCCGCCGCCAGGGCA---   364
            |
Db    3258  AGGTTTGGGGAAGAGACAACAACTCCCCTCCAGAAGCAGCGGCCGATAGACAAGGAACTG   3317

Qy     365  ---CCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCCCTCGTGTCCTGGCCATCAAGGTGG   421
            |
Db    3318  TATCCTTTAGCTTCCCCTCCCTGAACATCACTCTTTGGCAGCGGCCCCTCGTGTCACAATAAAGATAG   3377

Qy     422  GCGCCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCGCCGACGCTGTCTGGAGGAGA   481
            ||||
Db    3378  GGGCCCAGCTGTGAAGCTGCTGTGCGGCGCCGCCACCAGGCCGTGCTGCTGCTGGAGGAGA   3437

Qy     482  TGAGCCTGCCGGCCAAGTGGAAGCCCAAGATGATCGGCGCCATCGGCGGCTTCATCAAGG   541
            ||   ||||
Db    3438  TGAACCTCTCCCGGCCGTCTGGAGCCCAAGATGATCGGCGCCATCGGCGGCTTCATCAAGG   3497

Qy     542  TGCCCAGTAGGACCGATCAGGACGAGATCTGCGCGCAAGAGGCCATGGAGGGCCACCGTGC   601
            |||||
Db    3498  TGGGCCAGTAGCACGACGATCCTGATCGGGATCTGCGCGACCAAGGCCACCGGCACGCCCGTGC   3557

Qy     602  TGATCGGGCCCACCCCGGTGAACATCATCGGCCGCAACATGCTGACCCAGTTGGCTGCA   661
            ||
Db    3558  TGGTGGGCCCCACCCCGGTGAACATCATCGGCCGCAACCTGCTGACCCAGATCGGCTGCA   3617

Qy     662  CCCTGAACTTCCCCATCAGCCCCGTGGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGG   721
            ||   ||||
Db    3618  CCCTGAACTTCCCCATCATCAGCCCCGTGGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGG   3677

Qy     722  ACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCA   781
            |||||||
Db    3678  ACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGA   3737
```

```
Qy     782  TCTGGCGGAGGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCGAGAACCCCTACA   841
            |||||
Db    3738  TCTGCACCGGAGGATGGAGAAGGAGGGCAAGATCAGCAGGATCGGCCCGAGAACCCCTACA   3797

Qy     842  ACACCCCGTGTTCGCCATCAAGAGAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT   901
            ||||||||||
Db    3798  ACACCCCGTGTTCGCCATCAAGAGAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT   3857

Qy     902  TCCGGCAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC   961
            |||||||||||
Db    3858  TCCGGCAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACC   3917

Qy     962  CCGCCGGCCTGAAGAGAGAGAAGAGAGAGCGGCCAGCGGTGACGTGGGCGACGCCTACTTCA   1021
            |||||||||
Db    3918  CCGCCGGCCTGAAGCAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCA   3977

Qy    1022  GCGTGCCCTGGACGAGGACTTCCCAGTACACCGCCTTCACCATCCCCAGCATCAACA   1081
            ||||||||||||
Db    3978  GCGTGCCCTGGACAAGGACTTCCGGCAGCTCCAAGTACACCGCCTTCACCATCCCCAGCATCAACA   4037

Qy    1082  ACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGCTGAAGGGCCAGCC   1141
            ||||||||||
Db    4038  ACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGCTGAAGGGCCAGCC   4097

Qy    1142  CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCTTCCGGCCCCGCAACCCCG   1201
            ||
Db    4098  CCGCCATCTTCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCAACAGAACCCCCG   4157

Qy    1202  AGATCTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCC   1261
            ||||||||||
Db    4158  ACATCGGATCTTCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCC   4217

Qy    1262  AGCACCGCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCC   1321
            ||||||||||
Db    4218  AGCACCGCGCCACCAAGATCGAGGAGCTGCGCCAGCACCTGCTGCGCTGGGGCTTCACCACCC   4277

Qy    1322  CCGACAAGAGCACCAGAGAAGAGCCCCCCTCCGTGATGGGATGGGCTACGAGCTGCACCCCG   1381
            |||||||
Db    4278  CCGACAAGAGCACCAGAGAAGAGCCCCCCTCCGTGATGGGATGGGCTACGAGCTGCACCCCG   4337

Qy    1382  ACAAGTGACCGTGCAGCTGAACTGGGCGCCCATCGAGCTGCCCGAGAAGAGAGACCGGCATCAAGGTGC   1441
            ||
Db    4338  ACAAGTGGACCGTGCAGCTGAACTGGGCGCCCATCGAGCTGCCCGAGAAGAGACCGGCATCAAGGTGC   4397

Qy    1442  TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCCCAGGCGCCGAGATCTCTACCCGCGACATCGTGCCCCTGACCG   1501
            ||||||||||
Db    4398  TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCCCAGGCGCCGAGATCTACCCGCGACATCGTGCCCCTGACCG   4457

Qy    1502  GCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGAGGTGGTGCCCCTGACCG   1561
            |||||||||
Db    4458  GCCAGCTGTGCAAGCTGCTGCGCGGCGCCACCAAGGCCCTGACCGAGGTGGTGCCCCTGACCG   4517

Qy    1562  AGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAACCCGTGCAGGCGG   1621
            |||||||||
Db    4518  AGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGAAGGAGCCCGTGCAGGGCGG   4577

Qy    1622  TGTACTACGACCCCAGTAGGACCAAGGACCTGGTGGCCGAGAAGCAGCAGGCCACCAGT   1681
            ||
Db    4578  TGTACTACGACCCCACCAAGGACCTGATCGCGCAGACCTGATCGCGATCAGGGCCAGT   4637

Qy    1682  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACTGAAGACCCCCGCAAGA   1741
            |||
Db    4638  GGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACTGAAGACTGCCGCAAGA   4697

Qy    1742  TGTGCACCGCCCACACCAACAGCGTGAAGGTGCTGACAGAGGCCGTGCAGAAGATCGCCA   1801
            |
Db    4698  TGAAGGCGCCACCCACCAACCAACCAACGTGAAGGTGAAGCTGACAGAGGCCGTGCAGAAGATCGCCA   4757

Qy    1802  TGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGGCTGCCCATCCAGAAGGAGA   1861
            |
Db    4758  CCGAGAGCATCGTGATCTGGGGCAAGAGAGATCAAGTTCAAGCTGCCCATCCAGAAGGAGA   4817
```

Qy 131 GCAACTGCCGCGCCCCCGAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGAGGGCCACCAGA 190
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3087 GCAACTGCCGCGCCCCCGAAGAGAGGCTGCTGGAAGTGCGGCAAGGAGGAGGGCCACCAGA 3146

Qy 191 TGAAGGACTGACACCGAGCGCAGCAGGCTAA------------TAGGGAAGATCTGGCTTCCCACAAG 250
         |||||||||||||||||||||||||||
Db 3147 TGAAGGACTGACACCGAGCGCAGCAGGCTAA------TAGGGAAGATCTGGCTTCCCACAAG 3200

Qy 251 GCAAGGGCCCGCCGAGTTCCCAGCGAGGAACCGCGCAACAGCCCCACGCGCCGAGC 310
         |||||||||||||||||||||||||||||||
Db 3201 GGAAGGCCAGGGAATTTTCTTCAGAGCAGCAACCAGCAGCCAAGAGAGAGC 3260

Qy 311 CGACAACCCCCGCAGCAGGCCGGCCGCGGCCCAGGGCA 364
         TGCAGGTGCGCGG------
Db 3261 TTCAGGTTTGGGGAAGAGACAACAACTCCCTCAGAGCGCGCCCCTTGTGAGCATCAAGG 3320

Qy 365 CCCTGAACTTCCCAGATCACCCTGTGGCAGCGCCAGGAGGTGAAACATCGTGA 418

Db 3321 CTGTATCCTTTAGCTTCCTCAGATCACTCACTCGTTGGCAGCACCCCTCGTCGTCACAATAAAGA 3380

Qy 419 TGGGCGGCCAGATCAAGGAGCCCTGCTGGCCATCATCCAGGCCCGACAAGAGCGAGC 478
         |||||||||||||||||||||||||||||||
Db 3381 TAGGGGCGGCCAGCTGAAGGAGCCCTTCTGAGGATGCCCGACACCGGCCGTGCTGAGG 3440

Qy 479 AGATGAGCCTGGCCAGTGAAAGCCAAGTGAAGTGTACCTGAGCTGGGTGC 538
         ||||||||||||||||||||||||
Db 3441 AGATGAGCCTGGCCAGTGAAGCCAAGTGAAGTGTACCTGAGCTGGGTGC 3500

Qy 539 AGGTGCGCCAGATCGGCGCACGAGATCCTGATCGAAATCATCGGCCGCAACATGCTGACCCAGCTGGCT 598
         ||||||||||||||||||||||||||||
Db 3501 AGGTGGGCGTGACCAGATCCTGATCGAAATCATCGGCCGCAACATGCTGACCCAGATCGGCT 3560

Qy 599 TGCTGATCGGCCCCACCCCCGTGAAACATCAGCCCCATCAGCCCAT 658
         |||||||||||||||||||||||||||
Db 3561 TGCTGGTGGGCCCCACCCCCGTGAAACATCAGCCCCATCAGCCCAT 3620

Qy 659 GCACCCTGAACTTCCCCATCAGCCCCATCAGCCCAT 718
         ||||||||||||||||||||||||
Db 3621 GCACCCTGAACTTCCCCATCAGCCCCATCAGCCCAT 3680

Qy 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCCGACCGAGGAGAAGATCAAGGCCCTGACCG 778
         |||||||||||||||||||||||||||
Db 3681 TGGACGGCCCAAGGTGAAGCAGTGGCCCCGACCGAGGAGAAGATCAAGGCCCTGTGTGG 3740

Qy 779 CCATCTGCGAGGAGATGGAGAAGGAGGGCAAGTACCACGAAGATCGGCGCCCCGAGAACCCCT 838
         |||||||||||||||||||||||||||
Db 3741 AGATCTGCACCGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCGCCCCGAGAACCCCT 3800

Qy 839 ACAACACCCCCGTGTTCGCCATCAAGAAGACAGCAGCAAGTGGGCGCAAGCTGGTGG 898
         |||||||||||||||||||||||||||
Db 3801 ACAACACCCCCGTGTTCGCCATCAAGAAGACAGCAGCAAGTGGGCGCAAGCTGGTGG 3860

Qy 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCACTGGGCATCCCCC 958
         |||||||||||||||||||||||||||
Db 3861 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCACTGGGCATCCCCC 3920

Qy 959 ACCCCGCGCCTGAAGAAGACGAGAGAGTGGACCGTGACCGTGCTGCAGCGTGGGCGACGCCTACT 1018
         |||||||||||||||||||||||||||
Db 3921 ACCCCGCGCCTGAAGAAGACGAGAGAGTGGACCGTGACCGTGCTGCAGCGTGGGCGACGCCTACT 3980

Qy 1019 TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACCGGACTACACCGCCTTCACCACCCCAGCATCA 1078
         |||||||||||||||||||||||||||
Db 3981 TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACCGGACTACACCGCCTTCACCACCCCAGCATCA 4040

Qy 1079 ACAACGAGACCCCCGGCATCCGTTCTACCGTTCACGTCTCCAGGGCTGAAGGGCA 1138
         |||||||||||||||||||||||||||
Db 4041 ACAACGAGACCCCCGGCATCCGTTCTACCGTTCACGTCTCCAGGGCTGAAGGGCA 4100

Qy 1139 GCCCCAAGCATCTTCCAGAGCAGCATGAACCAAGATCCTGAGCC 1198
         |||||||||||||||||||||||||||
Db 4101 GCCCCAAGCATCTTCCAGTGCAGCATGAACCAAGATCCTGAGCC 4160

Qy 1199 CCGAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCGAGATCG 1258

Qy 1862 CCTGGGAGGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCG 1921
         |||||||||||||||||||||||||||
Db 4818 CCTGGGAGGAGGCCTGGTGGACCGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCG 4877

Qy 1922 TGAACACCCCCCCTGGTGAAGCTGTGGTACCAGTGGAGAAGGAGGCCATCATCGGCG 1981
         |||||||||||||||||||||||||||
Db 4878 TGAACACCCCCCCTGGTGAAGCTGTGGTACCAGTGGAGAAGGAGGCCATCATCGGCG 4937

Qy 1982 CCGAGACCTTCTACGTGGACGGCGCCAACCGCGAGATCGGCGAAGGCCGGCT 2041
         |||||||||||||||||||||||||||
Db 4938 CCGAGACCTTCTACGTGGACGGCGCCAACCGCGAGATCGGCGAAGGCCGGCT 4997

Qy 2042 ACGTGACCGACCGGGAGCCGGCCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAGA 2101
         |||||||||||||||||||||||||||
Db 4998 ACGTGACCGACCGGGAGCCGGCCGCGCAGAAGATCGTGGCCCCTGACCGAGACCACCAACCAGAGA 5057

Qy 2102 CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAAACATCGTGA 2161
         |||||||||||||||||||||||||||
Db 5058 CCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACAGCGGCCTGGAGGTGAACATCGTGA 5117

Qy 2162 CCGACAGCCAGTACGCCCTGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC 2221
         |||||||||||||||||||||||||||
Db 5118 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC 5177

Qy 2222 TGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGC 2281
         |||||||||||||||||||||||||||
Db 5178 TGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGC 5237

Qy 2282 CGCCCCACAAGGGCATCGGCGGCAACGAGATCGAGCAGCAGCTGGACGGCCTGGCAAGGCCATCC 2341
         |||||||||||||||||||||||||||
Db 5238 CCGCCCACAAGGGCATCGGCGGCAACGAGATCGAGCAGCAGCTGGACGGCCTGGCAAGGCCATCC 5297

Qy 2342 GCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCTGTGATCTACCAGTA 2393
         |||||||||||||||||||||||||||
Db 5298 GCAAGGTGCTGTTCCCGGACGGCATCGACGGCGGCATCGACGACACGAGAA 5349

RESULT 15
AX427921

LOCUS   AX427921   9189 bp   DNA   linear   PAT 20-JUN-2002
DEFINITION   Sequence 159 from Patent WO0232943.
ACCESSION   AX427921
VERSION   AX427921.1 GI:2153008
KEYWORDS   .
SOURCE   synthetic construct
  ORGANISM   synthetic construct
              artificial sequences.
REFERENCE   1
  AUTHORS   Huang,Y. and Nabel,G.J.
  TITLE   Modifications of hiv env, gag, and pol enhance immunogenicity for
              genetic immunization
  JOURNAL   Patent: WO 0232943-A 159 25-APR-2002;
              GOVERNMENT OF THE UNITED STATES (US)
FEATURES   Location/Qualifiers
     source   1..9189
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="plasmid pVR1012x/s containing HIV genes"
ORIGIN

Query Match   81.8%; Score 2018.6; DB 6; Length 9189;
Best Local Similarity   91.4%; Pred. No. 5.2e-203;
Matches 2190; Conservative 0; Mismatches 184; Indels 21; Gaps 4;

Qy 14   TGGCCGAGGCCATGAGCCAGG---CCACCACAGCGCCAACATCCTGATCGCAGCGCAGCAACT 70
         |||||||||||||||||||| ||||||
Db 2967 TGGCCCAGGGCCATGAGCCAGGTGACCAACAGCGCCATCATGATGCAGAGGGGCAACT 3026

Qy 71   TCAAAGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCC 130
         |||||||||||||||
Db 3027 TCCGCAACCAGCGCCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGGCCACACCGCCC 3086

Db 5241 TGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGCAGGTGGCGGCCTGGTGAGCGCCGGCA 5300
Qy 2339 TCCGCAAGGTGCTGTTCCTTGGACCGGCATCGATGGCGGCATCGTGATCTACCAGTA 2393
         ||||||||||||||||||||| || ||||| || | | |
Db 5301 TCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAGAGGAGCACGAGAA 5355

Search completed: April 10, 2004, 12:49:04
Job time : 6354.08 secs

Db 4161 CCGACAATCGTGATCTACCAGTACATGACGACGCCTGTACGTGGCCAGCGACCTGGAGATCG 4220
Qy 1259 GCCAGCACCGGCCAAGATCCAGAGAGCTGCCGCAAGCCACCTGCTGCGCCTGGGGCTTCACCA 1318
Db 4221 GCCAGCACCGCACCCAGAGATCGAGGAGCTGCCCAGCAGCTGCCGGCCACCTGGGGCTTCACCA 4280
Qy 1319 CCCCGACAAGAAGCACCAGAGAGGAGCCCCCTTCGTGGATGGGCTACGAGCTGCACC 1378
Db 4281 CCCCCGACAAGAAGCACCAGAGAGGAGCCCCCTTCGTGGATGGGCTACGAGCTGCACC 4340
Qy 1379 CCGACAAGTGGACCGTGCAGCCATCGAGCTGCCCGAGAAGGAGACTGGACCGTGAACG 1438
Db 4341 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACG 4400
Qy 1439 ACATCCAGAAGCTGGTGGCAAGCTGAACTGGCCAGATCTACCCCGGCATCAAGG 1498
Db 4401 ACATCCAGAAGCTGGTGGCAAGCTGAACTGGCCCCAGATCTACGCGGCATCAAGG 4460
Qy 1499 TGCGCCAGCTGTGTCAAGCTGCTGCGCGGCGGCCAAGGCGCCCTGACCGACATCGTGCCCCTGA 1558
Db 4461 TGCGCCAGCTGTGTCAAGCTGCTGCGCGGCCACCAAGGGCCCTGACCGACGTGGTGCCCCTGA 4520
Qy 1559 CCGAGGAGGCCGGCCAGCTGGAGAGCTGCCGAGAACCGCGGCGATCCTGCGGGAGCCGTGCACG 1618
Db 4521 CCGAGGAGGCCGGCCAGCTGGAGAGCTGCCGGAGAACCGCGGCGATCCTGAGGGAGCCGTGCACG 4580
Qy 1619 GCGTGTACTACGACACCCCAGCAAGGACCTGGTGCCGAGATCCAGAAGCAGGGGCCACGACC 1678
Db 4581 GCGTGTACTACGACACCCCAGCAAGGACCTGATCGCCGAGATCCAGAAGCAGGGGCCAGGGCC 4640
Qy 1679 AGTGGACCTACCAGATCTTACCCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGGCCA 1738
Db 4641 AGTGGACCTACCAGATCTTACCCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGCCC 4700
Qy 1739 AGATGCGCACCGCCCACCACCAACGACGTGAAGCAGCTGAAGCAGCCCGTGACCGAGAAGATCG 1798
Db 4701 GCATGAAGGGCGCGCCACCACCAACGACGTGAAGCAGCTGAAGCAGCCCGTGCAGAAGATCG 4760
Qy 1799 CCATGGAGAGCATCGTGATCTGGGGCCAAGACCCCCAGTTCCGCCTGCCCATCCCAGAAGG 1858
Db 4761 CCACCGAGAGCATCGTGATCTGGGGCCAAGACCCCCAAGTCAAGCTGCCCATCCCAGAAGG 4820
Qy 1859 AGACCTGGAGAGACCTGGTGGACCGACTACTGGGCCACCTGGATCCCCGAGTGGGGAGT 1918
Db 4821 AGACCTGGAGGAGACCTGGTGGTGGACCGACTACTGGGCCACCTGGATCCCCGAGTGGGGAGT 4880
Qy 1919 TCGTGAACACCCCCCTGGTGAAGCTTGTTACCAGTGACGCGGCCAACCGCGAGACCCATCATCG 1978
Db 4881 TCGTGAACACCCCCCCTGGTGAAGCTTGTTACCAGTGACGCGGCCAACCGCGAGACCCATCATCG 4940
Qy 1979 GCGCCGAGACCTTCTACGTGGACGGCGCGCCGCAGAAGATCGTGAGCTGCCAACCGCAAGGGCCG 2038
Db 4941 GCGCCGAGACCTTCTACGTGGACGGCGCGCCGCAGAAGATCGTGAGCTGCCAACCGCAAGGGCCG 5000
Qy 2039 GCTACGTGACCGACCGCGGGGGCGCCGGCCGCAGAAGATCGTGAGCCTGACCGAGACCAACAACAGA 2098
Db 5001 GCTACGTGACCGACCGCGGGGGGCGCCGGCCGCAGAAGGTGTGCCCCTGACCGACCACCAACAACAGA 5060
Qy 2099 AGACCGAGCTGCAGCCCATCCAGCTGGCCCCTGCAGGACAGCGGCAGCCGAGGTGAACATCG 2158
Db 5061 AGACCGAGCTGCAGCCCATCCACCTGGCCCCTGCAGGACAGCGGCAGCCGAGGTGAACATCG 5120
Qy 2159 TGACCGACACAGCCAGTCGGGCCTGGAAGCTGTGGTACCAGTGGAAGGAGAGCGGAGAGCG 2218
Db 5121 TGACCGACACAGCCAGTCGGGCCTGGAAGCTGTGGTACCAGTGGAAGGAGAGCGGAGAGCG 5180
Qy 2219 AGCTGGTGAACCAGATCATCGAGACAGATCATCGAGACAAGAAGGAGGAGGTGTACCTGAGCTGGG 2278
Db 5181 AGCTGGTGAGCCAGATCATCGAGACAGATCATCGAGACAAGAAGGAGGAGGTGTACCTGCCCTGGG 5240
Qy 2279 TGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGCAGCGTGTGAGCAAGGGCA 2338

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      April 10, 2004, 02:45:46 ; Search time 622.513 Seconds
                                         (without alignments)
                                       1849.133 Million cell updates/sec

Title:        US-09-610-313-30
Perfect score: 2469
Sequence:    1 gtcgacgccaccatggccga........gggctagcaccggtgaattc 2469

Scoring table:  IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:     3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:      6747726


Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq.29Jan04:*
        1:   geneseqn1980s:*
        2:   geneseqn1990s:*
        3:   geneseqn2000s:*
        4:   geneseqn2001as:*
        5:   geneseqn2001bs:*
        6:   geneseqn2002s:*
        7:   geneseqn2003as:*
        8:   geneseqn2003bs:*
        9:   geneseqn2003cs:*
       10:   geneseqn2004s:*


Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES
```

Result No.	Score	% Query Match Length	DB	ID	Description		
1	2469	100.0	2469	6	ABL39959	Abl39959	Synthetic
2	2457	99.5	2457	7	ACA03548	Aca03548	Synthetic
3	2457	99.5	2457	9	ADC13266	Adc13266	DNA of HI
4	2442.2	98.9	2463	6	ABL39960	Abl39960	Synthetic
5	2434.2	98.6	2457	7	ACA03547	Aca03547	Synthetic
6	2434.6	98.6	2457	9	ADC13265	Adc13265	DNA of HI
7	2415.4	97.8	2457	6	ABL39961	Abl39961	Synthetic
8	2401.8	97.3	2445	7	ACA03546	Aca03546	Synthetic
9	2401.8	97.3	2445	9	ADC13264	Adc13264	DNA of HI
10	2394.8	97.0	3930	9	ADC13230	Adc13230	DNA of HI
11	2393.2	96.9	3930	9	ADC13231	Adc13231	DNA of HI
12	2393.2	96.9	3930	9	ADC13232	Adc13232	DNA of HI
13	2393.2	96.9	5184	7	ACA03591	Aca03591	Synthetic
14	2393.2	96.9	5184	9	ADC13279	Adc13279	DNA of HI
15	2362.8	95.7	3531	9	ADC13234	Adc13234	DNA of HI
16	2361.2	95.6	3537	9	ADC13236	Adc13236	DNA of HI
17	2360.2	95.6	5145	7	ACA03521	Aca03521	Synthetic
18	2360.2	95.6	5145	9	ADC13233	Adc13233	DNA of HI
19	2350.2	95.2	3538	9	ADC13235	Adc13235	DNA of HI
20	2349.4	95.2	3624	7	ACA03550	Aca03550	Synthetic
21	2349.4	95.2	3624	9	ADC13268	Adc13268	DNA of HI
22	2301.8	93.2	3607	7	ACA03551	Aca03551	Synthetic
23	2301.8	93.2	3607	9	ADC13269	Adc13269	DNA of HI
24	2283.6	92.5	3597	7	ACA03549	Aca03549	Synthetic
25	2283.6	92.5	3597	9	ADC13267	Adc13267	DNA of HI
26	2149.6	87.1	2472	7	ACA03543	Aca03543	Synthetic
27	2149.6	87.1	2472	7	ACC78507	Acc78507	HIV p2Pol
28	2121.2	85.9	2466	7	ACA03542	Aca03542	Synthetic
29	2121.2	85.9	2466	7	ACC78506	Acc78506	HIV p2Pol
30	2094.4	84.8	2460	7	ACA03541	Aca03541	Synthetic
31	2094.4	84.8	2460	7	ACC78505	Acc78505	HIV p2Pol
32	2093.4	84.8	3564	7	ACA78488	Acc78488	HIV GagPo
33	2093.4	84.8	3564	7	ACC78489	Acc78489	HIV GagPo
34	2092.8	84.8	4716	7	ACA03522	Aca03522	Synthetic
35	2092.8	84.8	4716	9	ADC13238	Adc13238	DNA of HI
36	2089	84.6	3999	7	ACC78484	Acc78484	HIV GagCo
37	2087.4	84.5	3999	7	ACC78485	Acc78485	HIV GagCo
38	2087.4	84.5	3999	7	ACC78486	Acc78486	HIV GagCo
39	2087.4	84.5	5283	7	ACA03584	Aca03584	Synthetic
40	2087.4	84.5	5283	7	ACC78529	Acc78529	HIV TatRe
41	2087.2	84.5	4713	9	ACA03592	Aca03592	Synthetic
42	2087.2	84.5	4713	9	ADC13280	Adc13280	DNA of HI
43	2086.8	84.5	3462	9	ADC13237	Adc13237	DNA of HI
44	2061.4	83.5	3735	7	ACA03545	Aca03545	Synthetic
45	2061.4	83.5	3735	7	ACC78509	Acc78509	HIV p2Pol

```
                                            ALIGNMENTS

RESULT 1
ABL39959
ID   ABL39959 standard; DNA; 2469 BP.
XX
AC   ABL39959;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Synthetic construct PR975(+) SEQ ID NO:30.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO20204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Engelbrecht S,  Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 8; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
```

```
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 U; 0 Other;

Query Match             100.0%;  Score 2469;  DB 6;  Length 2469;
Best Local Similarity   100.0%;  Pred. No. 5.2e-298;
Matches 2469;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GTCGACGCCACCATGCGCGAGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GTCGACGCCACCATGCGCGAGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60

Qy      61 CGCAGCAACTTCAAGGGCCCCAAGGGCCATCATCAAGTGCTTCAACTGCGCCAAGGAGGGC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CGCAGCAACTTCAAGGGCCCCAAGGGCCATCATCAAGTGCTTCAACTGCGCCAAGGAGGGC 120

Qy     121 CACATCGCCCCCAACTGCCCGCCGCCCCGACGAGCGCCAGGCCAACTTCTTCCGGCAAGGAG 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CACATCGCCCCCAACTGCCCGCCGCCCCGACGAGCGCCAGGCCAACTTCTTCCGGCAAGGAG 180

Qy     181 GGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCCGGAGGACCTGGCC 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCCGGAGGACCTGGCC 240

Qy     241 TTCCCCAGGGCAAGGCCCGCCCAGTTCCCCAGCGAGCAGAACCGGCGAGGCGCCACC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TTCCCCAGGGCAAGGCCCGCCCAGTTCCCCAGCGAGCAGAACCGGCGAGGCGCCACC 300

Qy     301 AGCGCGGAGCTGCAGGTGCGGGCGACAACCCCGAGCGGAGGCGCGAGGCGCCAG 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGCGCGGAGCTGCAGGTGCGGGCGACAACCCCGAGCGGAGGCGCGAGGCGCCAG 360

Qy     361 GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTG 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTG 420

Qy     421 GGCGGGCCAGATCAAGGAGGAGCCCGCCGGCGCGCGACACACCGTGCCCCTGACC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGCGGGCCAGATCAAGGAGGAGCCCGCCGGCGCGCGACACACCGTGCCCCTGACC 480

Qy     481 ATGAGCTGCCCCGGCAAGTGGAAGCCCAGATGGCGGCATCGGCGGCTTCATCAAG 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ATGAGCTGCCCCGGCAAGTGGAAGCCCAGATGGCGGCATCGGCGGCTTCATCAAG 540

Qy     541 GTGTCGCCAGTACGACCCAGATCGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTG 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTGTCGCCAGTACGACCCAGATCGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTG 600

Qy     601 CTGATCGGCCCCACCCCCCGTGAACATCAGCCCGCCGCAGATCATCGCGGCCTGGGCTGC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CTGATCGGCCCCACCCCCCGTGAACATCAGCCCGCCGCAGATCATCGCGGCCTGGGCTGC 660

Qy     661 ACCCTGAACTTCCCCATCAGCCCGTGCCGCCATCAGCGAGAGCGTGAAGCTGGTGGAC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 ACCCTGAACTTCCCCATCAGCCCGTGCCGCCATCAGCGAGAGCGTGAAGCTGGTGGAC 720

Qy     721 GACGGGCCCAAGGTGAACGACGCCACCGAGCGCGACCGCACCGAGCGCGCCCTGACCGCC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GACGGGCCCAAGGTGAACGACGCCACCGAGCGCGACCGCACCGAGCGCGCCCTGACCGCC 780

Qy     781 ATCTGCGAGGAGATGAGAAGATGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ATCTGCGAGGAGATGAGAAGATGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840

Qy     841 AACACCCCGTGTTCGCCATCAAGGACAGCACCAAGCACCGTGGTGGAC 900
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AACACCCCGTGTTCGCCATCAAGAAGAGGACCACCAAGCACCGTGGTGGAC 900

Qy     901 TTCCGGCAGCTGAACAAGCGCACCCAGGACTTCTCGGGAGGTGCAGTCGAGTGCAGTCGCCATCCCCCAC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTCCGGCAGCTGAACAAGCGCACCCAGGACTTCTCGGGAGGTGCAGTCGAGTGCAGTCGCCATCCCCCAC 960

Qy     961 CCCGCCGGCCTGAAGAAGAAGAAGAGAGAGGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTC 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CCCGCCGGCCTGAAGAAGAAGAAGAGAGAGGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTC 1020

Qy    1021 AGCGTGCCCCTGGACGAGGACTTCCGGACAGTACACCGCCTTCACCATCCCAGCATCAAC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGCGTGCCCCTGGACGAGGACTTCCGGACAGTACACCGCCTTCACCATCCCAGCATCAAC 1080

Qy    1081 AACGAGACCCCCGGCATCGCTACCAGTACAACGTGCTGCCCAGGCTGGAAGGGCCAGC 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AACGAGACCCCCGGCATCGCTACCAGTACAACGTGCTGCCCAGGCTGGAAGGGCCAGC 1140

Qy    1141 CCCAGCATCTTCAGAGCAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCCAACCCCC 1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CCCAGCATCTTCAGAGCAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCCAACCCCC 1200

Qy    1201 GAGATCGTGATTTACCAGTACATGGACAGACCTGTACGGTGGGCAGCGAGCTGGAGATCGGC 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GAGATCGTGATTTACCAGTACATGGACAGACCTGTACGGTGGGCAGCGAGCTGGAGATCGGC 1260

Qy    1261 CAGCACCGGCCCAAGATGAGGAGCTGCGCAAGCACCTGCTGCGCGGCTGCTGGGGCTTCACCACC 1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CAGCACCGGCCCAAGATGAGGAGCTGCGCAAGCACCTGCTGCGCGGCTGCTGGGGCTTCACCACC 1320

Qy    1321 CCCGACAAGAGCACCAGAAGAGAGCCCCCCTTCGTGTGGATGGGCTACGAGCTGACCGAGGAC 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCCGACAAGAGCACCAGAAGAGAGCCCCCCTTCGTGTGGATGGGCTACGAGCTGACCGAGGAC 1380

Qy    1381 GACAAGTGACCGTGCAGCCATCGAGCTGCCCGAGAAGGAGCTGACCGTGAACGAC 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GACAAGTGACCGTGCAGCCATCGAGCTGCCCGAGAAGGAGCTGACCGTGAACGAC 1440

Qy    1441 ATCCAGAAGTGGTTGGGCAAGCTGAACTGGCGCCAAGATCTACCCCGGCATCAAGGTG 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 ATCCAGAAGTGGTTGGGCAAGCTGAACTGGCGCCAAGATCTACCCCGGCATCAAGGTG 1500

Qy    1501 CGCCAGCTGTGCAAGCTGCTGCGCGGCCAAGGCCCTGACCGACCATCGTGCCCCTGACC 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 CGCCAGCTGTGCAAGCTGCTGCGCGGCCAAGGCCCTGACCGACCATCGTGCCCCTGACC 1560

Qy    1561 GAGGAGGGCCAGTCAAGGAGGAGCTGGCGAGCTGGCCGAGATCCTGCGCGGAGGCCGTGCACGGC 1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GAGGAGGGCCAGTCAAGGAGGAGCTGGCGAGCTGGCCGAGATCCTGCGCGGAGGCCGTGCACGGC 1620

Qy    1621 GTGTACTACGACCCCAGCAAGATCCTGATCGAGATCCTGCGGCAAGCAAGGAGCAGGGCCACGACCAG 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GTGTACTACGACCCCAGCAAGATCCTGATCGAGATCCTGCGGCAAGCAAGGAGCAGGGCCACGACCAG 1680

Qy    1681 TGGACCTACCAGATCTACCACCACCCCAAGAACTCACCAGAGAGCCGGCAAGTACCGGCCAAG 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 TGGACCTACCAGATCTACCACCACCCCAAGAACTCACCAGAGAGCCGGCAAGTACCGGCCAAG 1740

Qy    1741 ATGGCACACGCCCCACCACACGTGAAGTGCTGGTTGGTACCAGCGTGAAGCAGTCGAAGATCGCC 1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 ATGGCACACGCCCCACCACACGTGAAGTGCTGGTTGGTACCAGCGTGAAGCAGTCGAAGATCGCC 1800

Qy    1801 ATGGAGAGCAGCAGTGGCCCCTGACCGAGGAGAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 ATGGAGGAGCATCGTGATCTGGGCAAGCACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1860

Qy    1861 ACCTGGGAGACCTGTGACCGACTACTGGCAGGCCGACTGGTCCCGAGAACCCCTAC 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 ACCTGGGAGACCTGTGACCGACTACTGGCAGGCCGACTGGTCCCGAGAACCCCTAC 1920

Qy    1921 GTGAACACCCCCCCCTGGTGAAGCTGCGTTGGTACCAGCGAAGGAGCCATCATCGGC 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 GTGAACACCCCCCCCTGGTGAAGCTGCGTGTGTACCAGCGAAGGAGCCCATCATCGGC 1980

Qy    1981 GCCGAGACCTTCTACGTGACGCGCACCCAGGCTGCAAGGCCGCCAAGGCCGGC 2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 GCCGAGACCTTCTACGTGACGCGCACCCAGGCTGCAAGGCCGCCAAGGCCGGC 2040
```

XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 2457 BP; 568 A; 830 C; 758 G; 301 T; 0 U; 0 Other;

Query Match   99.5%; Score 2457; DB 7; Length 2457;
Best Local Similarity 100.0%; Pred. No. 1.6e-296;
Matches 2457; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy         7 GCCACCATGCGCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGC   66
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GCCACCATGCGCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGC   60

Qy        67 AACTTCAAGGGCCCCAAGGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC  126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AACTTCAAGGGCCCCAAGGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC  120

Qy       127 GCCCGCAACTGCCGCGGCCCCCCCGCAAGGAAGGGCTGCTGGCAAGTGCGGCAAGGAGGGCCAC  186
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCCCGCAACTGCCGCGGCCCCCCCGCAAGGAAGGGCTGCTGGCAAGTGCGGCAAGGAGGGCCAC  180

Qy       187 CAGATGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC  246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CAGATGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC  240

Qy       247 CAGGGCAAGGCCCGAGTTCCCAGCGAGCAGAACCGGCCGAGGGCCGCCACCAGCCCGGC  306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CAGGGCAAGGCCCGAGTTCCCAGCGAGCAGAACCGGCCGAGGGCCGCCACCAGCCCGGC  300

Qy       307 GAGCTGCAGGTGCGGGGACAACCCCGCGCCGAGGGCCGCCGGGCGCCGAGGCGCCACC  366
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GAGCTGCAGGTGCGGGGACAACCCCGCGCCGAGGGCCGCCGGGCGCCGAGGCGCCACC  360

Qy       367 CTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGTGAGCATCAAGGTGGCGGGC  426
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCCTGTGAGCATCAAGGTGGCGGGC  420

Qy       427 CAGATCAAGGAGGACGCCTGCTGGACACCGCGGCCCCACGCACCGGTCTGAGGAGATGAGC  486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CAGATCAAGGAGGACGCCTGCTGGACACCGCGGCCCCACGCACCGGTCTGAGGAGATGAGC  480

Qy       487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAGGTGCGC  546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAGGTGCGC  540

Qy       547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCAAGAAGGCCATCGGCTGATC  606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCAAGAAGGCCATCGGCTGATC  600

Qy       607 GGCCCACCCCCCGTGAACATCATCGGCGCCGCCGCCAGCAGTGCTGACCCAGCTGGGCGTCGCACCCTG  666
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GGCCCACCCCCCGTGAACATCATCGGCGCCGCCGCCAGCAGTGCTGACCCAGCTGGGCGTCGCACCCTG  660

Qy       667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCCTGAAGCTGAAGCCCGGACGGCGGC  726
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCCTGAAGCTGAAGCCCGGACATGGACGGC  720

Qy       727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCATCTGC  786
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCATCTGC  780

Qy       787 GAGGAGGATGAGGAAGGAGGAGGGCCAAGATCACCAAGATCGGCGGCCCCGAGAACCCTACAACACC  846
```

RESULT 2
ACA03548
ID   ACA03548 standard; DNA; 2457 BP.
XX
AC   ACA03548;
XX
DT   22-MAY-2003 (first entry)
XX
DE   Synthetic DNA ecoding immunogenic HIV peptide #31.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO200304657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 36; 262pp; English.

```
Qy      2041 TACGTGACCGACCGGGGCCGCGCAGAGATGCTGAGCCTGACCGAGACCACCAGCGAG  2100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 TACGTGACCGACCGGGGCCGCGCAGAGATGCTGAGCCTGACCGAGACCACCAGCGAG  2100

Qy      2101 ACCGAGCTGCGAGGCCATCCAGCTGGCCCTGCCTGGCCTGCAGGACGCGGCAGCGGCCAGCGAGTGAACATCGTG  2160
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 ACCGAGCTGCGAGGCCATCCAGCTGGCCCTGCCTGGCCTGCAGGACGCGGCAGCGGCCAGCGAGGTGAACATCGTG  2160

Qy      2161 ACCGCAGCAGTAGCGCCCTGGGCATCATCCAGGCCCTGCCGAGGCCCAGGCCGACGAGAGCGGAG  2220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2161 ACCGCAGCCAGTACGCCCTGGGCATCATCCAGGCCCTGCCGAGGCCCAGGCCGACAAGAGCGGAGAGCGGAG  2220

Qy      2221 CTGGTGAACCAGATCATCGAGCAGCTGCACAACGAGGTGTACCTGAGCTGGGTG  2280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2221 CTGGTGAACCAGATCATCGAGCAGCTGCACAACGAGGTGTACCTGAGCTGGGTG  2280

Qy      2281 CCGCCCACAAGGGCATCGGCCGCAACGACAGCGCCATCGTGATCGTGCGCGATCGGCATCGGCAAGGGCATC  2340
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2281 CCGCCCACAAGGGCATCGGCCGCAACGACAGCGCCATCGTGATCGTGCGCGATCGGCATCGGCAAGGGCATC  2340

Qy      2341 CGCAAGGTGCTGTTCCTGGACGGCATCGGCCATCGGTGATCGTGATCCACCAGTACACATGGAC  2400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2341 CGCAAGGTGCTGTTCCTGGACGGCATCGGCCATCGGTGATCGTGATCCACCAGTACACATGGAC  2400

Qy      2401 GACCTGTACGTGGACGTGGCAGCGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC  2460
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2401 GACCTGTACGTGGACGTGGCAGCGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC  2460

Qy      2461 GGTGAATTC  2469
             |||||||||
Db      2461 GGTGAATTC  2469
```

```
Db   781 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  840
         GAGGAGGATGAGGAAGGAGGCCAAGATCACCAAGATCGGCGGCCCCGAGAACCCCTACACACC

Qy   847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC  906
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC  900

Qy   907 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCGCATCCCCCACCCGCC  966
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCGCC  960

Qy   967 GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG 1026
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG 1020

Qy  1027 CCCCTGGACGAGGACTTCCGCAAGTACACACCCTTCACCATCCCCAGCGTCAACAACGAG 1086
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CCCCTGGACGAGGACTTCCGCAAGTACACACCCTTCACCATCCCCAGCGTCAACAACGAG 1080

Qy  1087 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1146
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1140

Qy  1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATC 1206
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATC 1200

Qy  1207 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1266
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1260

Qy  1267 CGGCCAAGATCGAGGAGCTGCGCCAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1326
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CGGCCAAGATCGAGGAGCTGCGCCAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1320

Qy  1327 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1386
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1380

Qy  1387 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACACATCCAG 1446
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACACATCCAG 1440

Qy  1447 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAG 1506
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAG 1500

Qy  1507 CTGTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1566
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CTGTGCAAGCTGCTGCGCGGCACCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy  1567 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1626
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1620

Qy  1627 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACC 1686
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACC 1680

Qy  1687 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1746
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1740

Qy  1747 ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1806
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800

Qy  1807 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1866
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860

Qy  1867 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGTGAAC
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGTGAAC 1920

Qy  1927 ACCCCCCCCTGGTGAAGCTGTTGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1986
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ACCCCCCCCTGGTGAAGCTGTGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1980

Qy  1987 ACTTCTACGTGGACGGCCGCCAACCGGCAAGATCGGCCAAGGCCACCAACCAGAAGACCGAG 2046
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ACTTCTACGTGGACGGCCGCCAACCGGCAAGATCGGCCAAGGCCACCAACCAGAAGACCGAG 2040

Qy  2047 ACCGACCGGGGCCGCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2106
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ACCGACCGGGGCCGCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG 2100

Qy  2107 CTGCAGGCCATCAGCTGGCCCTGCCCGCCAGCAGACACGGCCAGCGAGTGAACATCGTGACCGAC 2166
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGCAGGCCATCAGCTGGCCCTGCCCGCCAGCAGACACGGCCAGCGAGTGAACATCGTGACCGAC 2160

Qy  2167 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACCAGAGAGCGAGAGCGAGCTGGTG 2226
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACCAGAGAGCGAGAGCGAGCTGGTG 2220

Qy  2227 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2286
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCCCC 2280

Qy  2287 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCGCAGGGCATCCCGCAAG 2346
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCGCAGGGCATCCGCAAG 2340

Qy  2347 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTG 2406
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTG 2400

Qy  2407 TACGTGGGCAGCGACGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2463
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 TACGTGGGCAGCGACGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
```

RESULT 3
ADC13266
ID   ADC13266 standard; DNA; 2457 BP.
XX
AC   ADC13266;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct p2Pol-opt_C SEQ ID NO 45.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,

PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines
XX
PS   Disclosure; Fig 42; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
XX   Type C related sequence of the invention.
SQ   Sequence 2457 BP; 568 A; 830 C; 758 G; 301 T; 0 U; 0 Other;

Query Match   99.5%;   Score 2457;   DB 9;   Length 2457;
Best Local Similarity   100.0%;   Pred. No. 1.6e-296;
Matches 2457;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      7 GCCACCATGGCCGAGGCCATGAGCGAGCCAGCGCCAGCGCCACAATCCTGATGCAGCGCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGAGGCCATGAGCGAGCCAGCGCCAGCGCCACAATCCTGATGCAGCGCAGC 60

Qy     67 AACTTCAAGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 120

Qy    127 GCCCGCAACTGCGGCCGCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCAC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCGCAACTGCGGCCGCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCAC 180

Qy    187 CAGATGAGGACTGCACCGAGCGGCCAGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAGGACTGCACCGAGCGGCCAGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 240

Qy    247 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACCCGCGCCAACAGCCCACCAGCCGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACCCGCGCCAACAGCCCACCAGCCGC 300

Qy    307 GAGCTGCAGGTGCGCGGCGACAACCCCCTGTGGGCAGCATCAAGGTGGGCGGC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGCTGCAGGTGCGCGGCGACAACCCCCTGTGGGCAGCATCAAGGTGGGCGGC 360

Qy    367 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCGCCGACACCGTGCTGGAGGAGATAGC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGGCCGCCGACACCGTGCTGGAGGAGATAGC 420

Qy    427 CAGATCAAGGAGGCCTGCTGGAAGTGCGGCAAGGAGGGCATCGGCCGTTCATCAAGGTGCGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGGCCTGCTGGAAGTGCGGCAAGGAGGGCATCGGCCGTTCATCAAGGTGCGC 480

Qy    487 CTGCCCGGCAAGTGAAGCCGGCATCGGCCGGCAAGAGAGGCCATCGGCCACCGTGCTGATC 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGCCCGGCAAGTGAAGCCGGCATCGGCCGGCAAGAGAGGCCATCGGCCACCGTGCTGATC 540

Qy    547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGCAGCCCTGAAGCTGACCCCTG 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGCAGCCCTGAAGCTGACCCCTG 600

Qy    607 GGCCCCACCCCCATCAGCCCGTGAACATGCTACCAGCTGGCCTGGCTGCCACCCTG 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCCCACCCCCATCAGCCCGTGAACATGCTACCAGCTGGCCTGGCTGCCACCCTG 660

Qy    667 AACTTCCCCATCAGCCCCATCAGCCCGTGAACGTGCCCGTGAAGCCCCGGCATGGACGCGGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AACTTCCCCATCAGCCCCATCAGCCCGTGAACGTGCCCGTGAAGCCCCGGCATGGACGCGGC 720

Qy    727 CCCAAGGTGAAGCAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCAAGGTGAAGCAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGC 780

Qy    787 GAGGAGATGGAGAAGGAGGAGAAGATCGGCCCCGAGAACCCTACCACACC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGATGGAGAAGGAGGAGAAGATCGGCCCCGAGAACCCTACCACACC 840

Qy    847 CCCGTGTTCGCCATCAAGAGAAGACAGCACAGCCAAGTGGCGGCAAGCTGGTGGACTTCCGC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCGTGTTCGCCATCAAGAGAAGACAGCACAGCCAAGTGGCGGCAAGCTGGTGGACTTCCGC 900

Qy    907 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 960

Qy    967 GGCCTGAAGAAGAAGAGAGCGCGTGACCGTGCTGGACGTGGCGACGCCTACTTCAGCGTG 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCTGAAGAAGAAGAGAGCGCGTGACCGTGCTGGACGTGGCGACGCCTACTTCAGCGTG 1020

Qy   1027 CCCCTGACGAGGACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCAACAACGAG 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGAGGACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCAACAACGAG 1080

Qy   1087 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1140

Qy   1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1200

Qy   1207 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGTGCGGCAAGGATCGGCCAGCAC 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGTGCGGCAAGGATCGGCCAGCAC 1260

Qy   1267 CGGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1320

Qy   1327 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGTGAACGAGATCCAG 1386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGTGAACGAGATCCAG 1380

Qy   1387 TGGACCGTGCAGCCCATCGAACTGCCCGAGAAGGACTCGGCCAGCATCAAGGTGCGCCAG 1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAACTGCCCGAGAAGGACTCGGCCAGCATCAAGGTGCGCCAG 1440

Qy   1447 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGGAG 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGGAG 1500

Qy   1507 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy   1567 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1620

Qy   1627 TACGACCCCAGCAAGGACCTGGTGGCCGAAGAGCAGCAGGGCCACGAGAGCAGTGGACC 1686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TACGACCCCAGCAAGGACCTGGTGGCCGAAGAGCAGCAGGGCCACGAGAGCAGTGGACC 1680

Qy   1687 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACTACGGCAAGATCGGCAAGATGCGC 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACTACGGCAAGATCGGCAAGATGCGC 1740

Qy   1747 ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAAGCTGCAGAAGATCGCCATGGAG 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAAGCTGCAGAAGATCGCCATGGAG 1800

Qy   1807 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG 1860
```

```
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 9; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 U; 0 Other;

Query Match            98.9%;  Score 2442.2;  DB 6;  Length 2463;
Best Local Similarity  99.6%;  Pred. No. 1.1e-294;
Matches 2460;  Conservative  0;  Mismatches 3;  Indels  3;  Gaps  1;

Qy     1  GTCGAGCGCCACCATGGCCGAGGCCATGAGCGCCCAGGCCGCCACCGGCCCAACATCCTGATGCAG   60
          ||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GTTGACGCGCCACCATGGCCGAGGCCATGAGCGCCCAGGCCGCCACCGGCCCAACATCCTGATGCAG   60

Qy    61  CGCAGCAACTTCAAGGGCCCCCAGAGCGCATCAAGTGCTTCAACTGCGGCAAGAGGGGC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CGCAGCAACTTCAAGGGCCCCCAGAGCGCATCAAGTGCTTCAACTGCGGCAAGAGGGGC  120

Qy   121  CACATCGCCCGCAACTGCCCGCCCCCCAAGAAGGGCTGCTGAAGTGCGGCAAGGAG  180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  CACATCGCCCGCAACTGCCCGCCCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG  180

Qy   181  GGCCACCAGATGAAGGACTGCACCGAGCGCCAAGCTTCTTCCGCGAGGACCTGGCC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  GGCCACCAGATGAAGGACTGCACCGAGCGCCAAGCTTCTTCCGCGAGGACCTGGCC  240

Qy   241  TTCCCCCAGGGCAAGGCCCGAGTTCCCCAGCGAGCGCCGAGCAGCAGCCCCACC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TTCCCCCAGGGCAAGGCCCGAGTTCCCCAGCGAGCGCCGAGCAGCAGCCCCACC  300

Qy   301  AGCCGCGAGCTGCCAGTGCGCGGGCGACAACCCGCGAGGCCGGCCGAGCGCCAG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  AGCCGCGAGCTGCCAGTGCGCGGGCGACAACCCGCGAGGCCGGCCGAGCGCCAG  360

Qy   361  GGCACCCTGAACTTCCCCAGATCACCCGTGTGGCACCATCAACGGCATCAAGGTG  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  GGCACCCTGAACTTCCCCAGATCACCCGTGTGGCACCATCAACGGCATCAAGGTG  420

Qy   421  GGCGGCCAGATCAAGAAGCCCTGCTGGACACCGGCCGCGACACCGGCATCAAGGAG  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  GGCGGCCAGATCAAGAAGCCCTGCTGGACACCGGCCGCGACACCGGCATCAAGGAG  480

Qy   481  ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGGCTTCATCAAG  540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481  ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGGCTTCATCAAG  540

Qy   541  GTGCGGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCATCGGCACCGTG  600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541  GTGCGGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCATCGGCACCGTG  600

Qy   601  CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCAACATCATGTGAACATCGTGGGCTGC  660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCAACATCATGTGAACATCGTGGGCTGC  660

Qy   661  ACCCTGAACTTCCCATCAGCAGCCCCATCAGCAGCCCGTGAAGCTGAAGCCCGCATG  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  ACCCTGAACTTCCCATCAGCAGCCCCATCAGCAGCCCGTGAAGCTGAAGCCCGGCATG  720
```

```
Qy  1867  GAGACCTGCTGGACCGACTACTGCCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGTGAAC  1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  GAGACCTGCTGGACCGACTACTGCCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGTGAAC  1920

Qy  1927  ACCCCCCCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCATCGGCGGCCGAG  1986
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTCGAGAAGGAGCCCATCATCGGCGGCCGAG  1980

Qy  1987  ACCTTCTACGTGGACAGGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG  2046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  ACCTTCTACGTGGACAGGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTG  2040

Qy  2047  ACCGACCGGGGCCCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCCAG  2106
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  ACCGACCGGGGCCCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCCAG  2100

Qy  2107  CTGCAGGCCATCCAGCTGCCCTGCAGGACAGCAGGCCAGCGAGGTGAACATCGTGACCGAC  2166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  CTGCAGGCCATCCAGCTGCCCTGCAGGACAGCAGGCCAGCGAGGTGAACATCGTGACCGAC  2160

Qy  2167  AGCCAGTAGCCCCTGGGCATCATCCAGGCCCAGCCCGACAAAGAGCGAGGAGGCAGCTGTG  2226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  AGCCAGTAGCCCCTGGGCATCATCCAGGCCCAGCCCGACAAAGAGCGAGGAGGCAGCTGTG  2220

Qy  2227  AACCAGATCATCGAGCAGCTGATCAAGAGAGGAGAAGTGTACCTGAGCTGGGTGCCGCC  2286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  AACCAGATCATCGAGCAGCTGATCAAGAGAGGAGAAGTGTACCTGAGCTGGGTGCCGCC  2280

Qy  2287  CACAAGGGCATCGGCGGCCAACGAGCAGATCGAACAGAGTCGGTGAGCAAGGGCATCCGCAAG  2346
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CACAAGGGCATCGGCGGCCAACGAGCAGATCGAACAGAGTCGGTGAGCAAGGGCATCCGCAAG  2340

Qy  2347  GTGCTGTTCCTGGACGGCATCGATGCGGCGCATCGATCCTGATCCTGATCCTGGACGACCTG  2406
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GTGCTGTTCCTGGACGGCATCGATGCGGCGCATCGATCCTGATCCTGATCCTGGACGACCTG  2400

Qy  2407  TACGTGGGCAGCGGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2463
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  TACGTGGGCAGCGGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2457
```

```
RESULT 4
ABL39960

ID   ABL39960 standard; DNA; 2463 BP.
XX
AC   ABL39960;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic construct PR975YM SEQ ID NO:31.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO020204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.
```

Left column (alignment):

```
Qy   721  GACGGCCCCAAGGTGAAGCAGTGCCCTGACCGAGGAGAAGATCAGGCCCTGACCGCC  780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  GACGGCCCCAAGGTGAAGCAGTGCCCTGACCGAGGAGAAGATCAGGCCCTGACCGCC  780

Qy   781  ATCTGCGAGGAGATGAGAGAAGGGCCAAGCAAGATCACCACCGAACCCCTAC  840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  ATCTGCGAGGAGATGAGAGAAGGGCCAAGCAAGATCACCACCGAACCCCTAC  840

Qy   841  AACACCCCGTGTTCGCCATCAAGAGGACAGCACCCAAGTGGCGCAAGCTGGTGGAC  900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  AACACCCCGTGTTCGCCATCAAGAGGACAGCACCCAAGTGGCGCAAGCTGGTGGAC  900

Qy   901  TTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCAC  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  TTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCCAC  960

Qy   961  CCCGGCCGGCCTGAAGAGAGAAGAGCGTGACGGTGGCGACGCGCTACTTC  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  CCCGGCCGGCCTGAAGAGAGAAGAGCGTGACGGTGGCGACGCGCTACTTC  1020

Qy  1021  AGCGTGCCCCTGACGAGGACTTCGGCCAAGTACACCGCCTTCACCATCCCCAGCATCAAC  1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  AGCGTGCCCCTGACGAGGACTTCGGCCAAGTACACCGCCTTCACCATCCCCAGCATCAAC  1080

Qy  1081  AACGAGACCCCCGGCATCCGCTACCAGTACACACGCTGCTGCCCCAGGGCTGGAAGGGCAGC  1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  AACGAGACCCCCGGCATCCGCTACCAGTACACACGCTGCTGCCCCAGGGCTGGAAGGGCAGC  1140

Qy  1141  CCCAGCATCTTCCAGGAGCAGCATGACCAAGATCCTGGAGCCCTTCGGCGCGCCGCAACCCC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CCCAGCATCTTCCAGGAGCAGCATGACCAAGATCCTGGAGCCCTTCGGCGCGCCGCAACCCC  1200

Qy  1201  GAGATCGTGATCTTACCAGTCACA-----GGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGC  1260
          ||||||||||||||||||         ||||||||||||||||||||||||||||||||||
Db  1201  GAGATCGTGATCTTACCAGTCACGTGTACGTGGGCAGCGACCTGGAGATCGGC  1254

Qy  1261  CAGCACCCGCCAAGATCGAGAGATCGAGAGCTGCTCGCCTCGGGCTTCACCACC  1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255  CAGCACCCGCCAAGATCGAGAGATCGAGAGCTGCTCGCCTCGGGCTTCACCACC  1314

Qy  1321  CCCGACAAGAGCACCAGAAGAGAGGAGGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCGACCCC  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315  CCCGACAAGAGCACCAGAAGAGAGGAGGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCGACCCC  1374

Qy  1381  GACAAGTGACCGTGCAGCCATCGAGCTGCCCCGAGGAGAGAGCTGACCGTGAACGAC  1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1375  GACAAGTGACCGTGCAGCCATCGAGCTGCCCCGAGGAGAGAGCTGACCGTGAACGAC  1434

Qy  1441  ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCCAGCAGCTGCCCGGCATCAAGGTG  1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1435  ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCCAGCAGCTGCCCGGCATCAAGGTG  1494

Qy  1501  CGCCAGCTGTGCAAGCTGCTCGCGCGGCGCAAGGCCCGCCAAGATCGTGCCCCTGACC  1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1495  CGCCAGCTGTGCAAGCTGCTCGCGCGGCGCAAGGCCCGCCAAGATCGTGCCCCTGACC  1554

Qy  1561  GAGGAGGGCCAGCTGGAGCTGGCCCGAGAACCGGCCGAGATCCTGCGCGAGCCCGTGCACGGC  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1555  GAGGAGGGCCAGCTGGAGCTGGCCCGAGAACCGGCCGAGATCCTGCGCGAGCCCGTGCACGGC  1614

Qy  1621  GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGGACCAG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1615  GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGGACCAG  1674

Qy  1681  TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG  1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1675  TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG  1734

Qy  1741  ATGCGCACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCC  1800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1735  ATGCGCACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCC  1794

Qy  1801  ATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAGGAG  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Right column (alignment):

```
Db  1795  ATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCGAGAGGAG  1854

Qy  1861  ACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGGAGTTC  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1855  ACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGGAGTTC  1914

Qy  1921  GTGAACAACCCCCCCCCCTGGTGAAGCTGGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1915  GTGAACAACCCCCCCCCCTGGTGAAGCTGGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC  1974

Qy  1981  GCCGAGACCTTCACGTGACGGCCCGCCAACCGGACAAGATCGGCACACCAGGCCGGC  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1975  GCCGAGACCTTCACGTGACGGCCCGCCAACCGGACAAGATCGGCACACCAGGCCGGC  2034

Qy  2041  TACGTGACGACCGGGGCCGGCAGAAGACGTGAGCGCCGAGGTGACAATCGTG  2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2035  TACGTGACGACCGGGGCCGGCAGAAGACGTGAGCGCCGAGGTGACAATCGTG  2094

Qy  2101  ACCGAGCTGCAGGCCATCGAGCTGCCCTCCAGGACAGCGGCAGCTGACATCGTG  2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2095  ACCGAGCTGCAGGCCATCGAGCTGCCCTCCAGGACAGCGGCAGCTGAACATCGTG  2154

Qy  2161  ACGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGACAAGAGGAGCGAG  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2155  ACGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGACAAGAGGAGCGAG  2214

Qy  2221  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG  2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2215  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG  2274

Qy  2281  CCCGCCCACAAGGGATCGGCGCGGCATCGATGACGAGCAGATCGAACAAGCTGGTGAGCAAGGGCATC  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2275  CCCGCCCACAAGGGATCGGCGCGGCATCGATGACGAGCAGATCGAACAAGCTGGTGAGCAAGGGCATC  2334

Qy  2341  CGGCAAGGTGCTGTTCCTGGACGCGGGCGGCATCGATGGCGCGCAGCCTAGCAGTACATGGAC  2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2335  CGGCAAGGTGCTGTTCCTGGACGCGGGCGGCATCGATGGCGCGCAGCCTAGCAGTACATGGAC  2394

Qy  2401  GACCTGTACGTGGGCAGCGGCGGCGGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC  2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2395  GACCTGTACGTGGGCAGCGGCGGCGGCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACC  2454

Qy  2461  GGTGAATTC  2469
          |||||||||
Db  2455  GGTGAATTC  2463
```

```
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 35; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptide. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 2457 BP; 564 A; 835 C; 758 G; 300 T; 0 U; 0 Other;

Query Match             98.6%;  Score 2434.6;  DB 7;  Length 2457;
Best Local Similarity   99.6%;  Pred. No. 9.7e-294;
Matches 2453;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps  1;
```

```
Qy      1 GTCGACGCCACCATGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCGACGCCACCATGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATGCAG 60

Qy     61 CGCAGCAACTTCAAGGGCCCCAAGGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCAGCAACTTCAAGGGCCCCAAGGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGG 120

Qy    121 CACATCGCCCGACAACTGCCCGCGCGCCCCCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACATCGCCCGACAACTGCCCGCGCGCCCCCCCCGGCAAGAAGGGCTGCTGGAAGTGCGGC 180

Qy    181 GGCCACCAGATGAAGGACTGCACCGAGTGCGGCGCGGCCAGGCCAACTTCTTCCGCGAGCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCCACCAGATGAAGGACTGCACCGAGTGCGGCGCGGCCAGGCCAACTTCTTCCGCGAGCC 240

Qy    241 TTCCCCAGGGCAGCGGCAAGGCCGCGGCGAGTTCCCCAGCGAGCAGAACCGCGCAACAGCCCCACC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCCCCAGGGCAGCGGCAAGGCCGCGGCGAGTTCCCCAGCGAGCAGAACCGCGCAACAGCCCCACC 300

Qy    301 AGCCGCGAGTCGCAGGTGCGGCGGCGGCGAGTTCCCGGACACCCCGGAGGCGGCGGCCGAGGCGCCAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCCGCGAGTCGCAGGTGCGGCGGCGGCGAGTTCCCGGACACCCCGGAGGCGGCGGCCGAGGCGCCAG 360

Qy    361 GGCACCCTGAACTTCCCCGAGCAAGGAGGAGCCCTGTGCAGCACCGGACACCGCTGGCATCAAGGTG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCGCACCCTGAACTTCCCCCAGATCGGCAGCCCTGTGCGGCACCGAGCACCGCTGGCATCAAGGTG 420

Qy    481 ATGAGCCTGCCCGGCAAGTGGAAGCCGACAAGATGATCGGCCGGCTTCATCAAG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGAGCCTGCCCGGCAAGTGGAAGCCGACAAGATGATCGGCCGGCTTCATCAAG 540

Qy    541 GTGCGCCGAGTACGACCAGATCCTGATCGAGATGAGAAGGCCATCGGCACGCCGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGCGCCGAGTACGACCAGATCCTGATCGAGATGAGAAGGCCATCGGCACGCCGTG 600
```

```
Qy    541 GTGCGCCGAGTACGACCAGATCCTGATGAGATCATCGGCGCAACATCATCGGCGCACAGAAGGCCATCGGCACGCCGTG 600
Db    601 CTGATCGGCCCCACCCCGTGCCCGCCGTGAACATCATCGGCGCCACATGCTGACCCAGCTGGGCTGC 660
```

```
Qy    601 CTGATCGGCCCCACCCCGTGCCCGCCGTGAACATCATCGGCGCCACATGCTGACCCAGCTGGGCTGC 660
Db    601 CTGATCGGCCCCACCCCGTGCCCGCCGTGAACATCATCGGCGCCACATGCTGACCCAGCTGGGCTGC 660
```

```
Qy    661 ACCCTGAACTTCCCATCAGCCCATCGAGACCCCATCAGCCCCATCAGCCCCGTGAAGCTGAAGCCCGGCATG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ACCCTGAACTTCCCATCAGCCCATCGAGACCCCATCAGCCCCATCAGCCCCGTGAAGCTGAAGCCCGGCATG 720
```

```
Qy    721 GACGGCCCCAAGGTGAAGCAGTGCGCCCTGACCGAGAGAGATCAAGGCCTACCGCC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACGGCCCCAAGGTGAAGCAGTGCGCCCTGACCGAGAGAGATCAAGGCCCTGACCGCC 780
```

```
Qy    781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGCCCCGAGAACCCCTAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGCCCCGAGAACCCCTAC 840
```

```
Qy    841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
```

```
Qy    901 TTCCGGGAGTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTCCGGGAGTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
```

```
Qy    961 CCGGCCGGCCGCTGAGAAGAAGAAGAGCGTGACCGTGGGCGACGCGCCTACTTC 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CCGGCCGGCCGCTGAGAAGAAGAAGAGCGTGACCGTGGGCGACGCGCCTACTTC 1020
```

```
Qy   1021 AGGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACGCGCCTTCACCATCCCCAGCATCAAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGCCGTGCCCCTGGACGAGGACTTCCGCAAGTACACGCGCCTTCACCATCCCCAGCATCAAC 1080
```

```
Qy   1081 AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
```

```
Qy   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC 1200
```

```
Qy   1201 GAGATGCTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGC 1254
          |||||||||||||||||||||||||||||      ||||||||||||||||||||||||
Db   1201 GAGATGCTGATCTACCA------GGCCCCCCCGTACGTGGGCAGCGACCTGGAGATCGGC 1260
```

```
Qy   1261 CAGCACCGGCCAAGATCGAGGAGTGCGGCAAGCACCTGCTGCGCTGGGGCTTCACCACC 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 CAGCACCGGCCAAGATCGAGGAGTGCGGCAAGCACCTGCTGCGCTGGGGCTTCACCACC 1314
```

```
Qy   1321 CCGACAAGAGCACCAGAGGACACCACCAGAAGAGAGCTACCCCGGGATCGTGCCACCCC 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 CCGACAAGAGCACCAGAGGACACCACCAGAAGAGAGCTACCCCGGGATCGTGCCACCCC 1374
```

```
Qy   1381 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGAC 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACGAC 1434
```

```
Qy   1441 ATCCAGAAGCTGGTGGGCAAGCTGGTCTGCGGCGGGCCCGGCATCGTGCCCCTGACC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1435 ATCCAGAAGCTGGTGGGCAAGCTGGTCTGCGGCGGGCCCGGCATCGTGCCCCTGACC 1494
```

```
Qy   1501 GGCGAGCTGTGCAAGCTGCTGCGCGGAGTCGAGCTGGCCAAGGCCCGAGAACCGCGGC 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1495 CGCAGCTGTGCAAGCTGCTGCGCGGAGTCGAGCTGGCCAAGGCCCGAGAACATCGTGCCCTGACC 1554
```

```
Qy   1561 GAGGAGGCCGAGCTGGAGATCTGCGGCGGCTTCATCAAG 1614
          |||||||||||||||||||||||||||||||||||||||
Db   1555 GAGGAGGCCGAGCTGGAGATCTGCGGCGGCTTCATCAAG 1614
```

```
Qy   1621 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGAGGCCATCGGCACGACCAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1615 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGAGGCCATCGGCACGACCAG 1674
```

PN　WO200304620-A2.
XX
PD　16-JAN-2003.
XX
PF　05-JUL-2002;　2002WO-US021420.
XX
PR　05-JUL-2001;　2001US-0303192P.
PR　31-AUG-2001;　2001US-0316860P.
PR　16-JAN-2002;　2002US-0349871P.
XX
PA　(CHIR) CHIRON CORP.
PA　(UYST-) UNIV STELLENBOSCH.
XX
PI　Zur Megede J,　Barnett SW,　Lian Y,　Engelbrecht S,　Van Rensburg EJ;
XX
DR　WPI; 2003-221593/21.
XX
PT　New expression cassette comprising a polynucleotide sequence encoding a
PT　polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT　Prot, or Rev polypeptide, useful for immunization, or generating
PT　packaging cell lines.
XX
PS　Disclosure; Fig 41; 301pp; English.
XX
CC　The invention relates to a novel expression cassette comprising a
CC　polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC　Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC　expression cassette can be used to treat HIV type C by gene therapy or
CC　used in the development of a vaccine. The gene delivery vector is
CC　administered intramuscularly, intramucosally, intranasally,
CC　subcutaneously, intradermally, transdermally, intravaginally,
CC　intrarectally, orally or intravenously. The expression cassette is useful
CC　for immunisation, generating packaging cell lines and producing HIV
CC　polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC　Type C related sequence of the invention.
XX
SQ　Sequence 2457 BP; 564 A; 835 C; 758 G; 300 T; 0 U; 0 Other;

Query Match　　98.6%;　Score 2434.6;　DB 9;　Length 2457;
Best Local Similarity　99.6%;　Pred. No. 9.7e-294;
Matches 2453;　Conservative　0;　Mismatches　4;　Indels　6;　Gaps　1;

Qy　　1　GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGGGCCACCAACATCCTGATGCAG　60
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　1　GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGGGCCACCAACATCCTGATGCAG　60

Qy　61　CGCAGCAACTTCAAGGGCCCCAAGGGCATCATCAAGTGCTTCAACTGCGGCAAGAGGGGC　120
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　61　CGCAGCAACTTCAAGGGCCCCAAGGGCATCATCAAGTGCTTCAACTGCGGCAAGAGGGGC　120

Qy　121　CACATCGCCCCAACTGCGGCCCCCCCGCCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG　180
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　121　CACATCGCCCCAACTGCGGCCCCCCCGCCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG　180

Qy　181　GGCCACCAGATGAAGGACTGCACCGAGCGGCAGGCCCAACTTCTTCCGGAGGACCTGGCC　240
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　181　GGCCACCAGATGAAGGACTGCACCGAGCGGCAGGCCCAACTTCTTCCGGAGGACCTGGCC　240

Qy　241　TTCCCCAGGGCAAGGCCCGGAGTTCCCAGCGAGCAGCAGAACCGGCCAACAGCCCCACC　300
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　241　TTCCCCAGGGCAAGGCCCGGAGTTCCCAGCGAGCAGCAGAACCGGCCAACAGCCCCACC　300

Qy　301　AGCCGCGGAGCTGCAGGTGCGGCGGCGACAACCCCGGCCGCGGAGGCCCGGAGCCGCCAG　360
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　301　AGCCGCGGAGCTGCAGGTGCGGCGGCGACAACCCCGGCCGCGGAGGCCCGGAGCCGCCAG　360

Qy　361　GGCACCCTGAACTTCCCCAGATCACCTGGCGCCCCTGTGGCAGCATCAAGGTG　420
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　361　GGCACCCTGAACTTCCCCAGATCACCTGGCGCCCCTGTGGCAGCATCAAGGTG　420

Qy　421　GGCGGCCAGATCAAGGAGCCCCTGCTGCCGACACCGCGGCCGCCGACACCGCGTGCTGGAGGAG　480
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　421　GGCGGCCAGATCAAGGAGCCCCTGCTGCCGACACCGCGGCCGCCGACACCGCGTGCTGGAGGAG　480

Qy　1681　TGGACCTACCAGATTACCAGGAGGCCCTTCAAGAACCTGAAGACCTGAAGACCGGCAAGTACGCCAAG　1740
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1675　TGGACCTACCAGATCACCAGGAGGCCCTTCAAGAACCTGAAGACCTGAAGACCGGCAAGTACGCCAAG　1734

Qy　1741　ATGCGCACCGCCCACACCAACCACACCAACGACGTGAAGCAGCTGACCGAGCGGATCGCC　1800
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1735　ATGCGCACCGCCCACACCAACCACACCAACGACGTGAAGCAGCTGACCGAGCGGATCGCC　1794

Qy　1801　ATGGAGAGCATCGTGATCGTGGGCAAGACCCCCAGTTCCGCCTGCCCATCCAGAAGGAG　1860
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1795　ATGGAGAGCATCGTGATCGTGGGCAAGACCCCCAGTTCCGCCTGCCCATCCAGAAGGAG　1854

Qy　1861　ACCTGGGAGACCTGGTGGACCGACTACTAGGCCACCTGGATCCCGAGTGGGAGTTC　1920
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1855　ACCTGGGAGACCTGGTGGACCGACTACTAGGCCACCTGGATCCCGAGTGGGAGTTC　1914

Qy　1921　GTGAACAACCCCCCCCCGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC　1980
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1915　GTGAACACCCCCCCCCCGTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC　1974

Qy　1981　GCCGAGACCTTCTACGTGGACGGGCGCGCCGCCAACCCGCGAGACCAAGATCGGCAAGGCCGGC　2040
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　1975　GCCGAGACCTTCTACGTGGACGGGCGCGCCGCCAACCCGCGAGACCAAGATCGGCAAGGCCGGC　2034

Qy　2041　TACGTGACCGACCGGCGGGGCGCCGCAGAGATCGTGAGCCTGACCGAGACGCTGCTCAACCAGAAG　2100
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2035　TACGTGACCGACCGGCGGGGCGCCGCAGAGATCGTGAGCCTGACCGAGACGCTGCTCAACCAGAAG　2094

Qy　2101　ACCGAGCTGCAGGCCATCCAGTGCCCTGCCAGACAGCGGCGAGGTGAACATCGTG　2160
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2095　ACCGAGCTGCAGGCCATCCAGTGCCCTGCCAGACAGCGGCGAGGTGAACATCGTG　2154

Qy　2161　ACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCAGACAAGAGCGGAGAGCGAG　2220
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2155　ACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCAGACAAGAGCGGAGAGCGAG　2214

Qy　2221　CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG　2280
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2215　CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG　2274

Qy　2281　CCCGCCCACAAAGGGCATCGGCGCCAACGAGCAGATGACAAGCTGCTGTGAGCAAGGGCATC　2340
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2275　CCCGCCCACAAGGGCATCGGCGCCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATC　2334

Qy　2341　CGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGAC　2400
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2335　CGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGAC　2394

Qy　2401　GACCTGTACGTGGGCAGCGGCGGCCGCGCCCTAGGATCGATTAAAAGCTTCCGGGGGCTAGCACC　2460
　　　　　|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　2395　GACCTGTACGTGGGCAGCGGCGGCCGCGCCCTAGGATCGATTAAAAGCTTCCGGGGGCTAGCACC　2454

Qy　2461　GGT　2463
　　　　　|||
Db　2455　GGT　2457

RESULT 6
ADC13265
ID　ADC13265 standard; DNA; 2457 BP.
XX
AC　ADC13265;
XX
DT　18-DEC-2003　(first entry)
XX
DE　DNA of HIV construct p2Pol-opt-YM_C SEQ ID NO 44.
XX
KW　expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW　Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS　Human immunodeficiency virus.
XX

Qy   1561 GAGGAGGCCGAGCTGGAGCTGCCGAGAACCGCGAGATCCTGCGCGGCAGCCCGTGCGCGGC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1555 GAGGAGGCCGAGCTGGAGCTGCCGAGAACCGCGAGATCCTGCGCGGCAGCCCGTGCGCGGC 1614

Qy   1621 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCAGCAGGGCCACGACCAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1615 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCAGCAGGGCCACGACCAAG 1674

Qy   1681 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAGACCGGCAAGTACGCCAAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1675 TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAGAACCGGCAAGTACGCCAAG 1734

Qy   1741 ATGCGCACCGCCCACACCAACGACGTGAAGCAGCTGAACAGCTGACGGCCGTGCAGAAGATCGCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1735 ATGCGCACCGCCCACACCAACGACGTGAAGCAGCTGAACAGCTGACGGCCGTGCAGAAGATCGCC 1794

Qy   1801 ATGGAGAGCATCGTGATCGGGCAAGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1795 ATGGAGAGCATCGTGATCGGGCAAGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1854

Qy   1861 ACCTGGGAGACTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1855 ACCTGGGAGACTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC 1914

Qy   1921 GTGAACACCCCCCCCTGGTGAAGCTGGTGGTTACCAGCGTGGAGGAGCCCAAGGGCCGGC 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1915 GTGAACACCCCCCCCTGGTGAAGCTGGTGGTTACCAGCGTGGAGGAGCCCAAGGGCCGGC 1974

Qy   1981 GCCGAGACCTTCTACGTGGACGGCGCCGGCCCAACCGCGAGACCAAGATCGGCAAGGGCCGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1975 GCCGAGACCTTCTACGTGGACGGCGCCGGCCCAACCGCGAGACCAAGATCGGCAAGGGCCGC 2034

Qy   2041 TACGTGACCGACCGGCGCCGGCAGAAGATCGTGAGCTGACCGAGACCACCAGAAG 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2035 TACGTGACCGACCGGCGCCGGCAGAAGATCGTGAGCTGACCGAGACCACCAGAAG 2094

Qy   2101 ACCGAGCTGCAGGCCATCCAGTCGCAGGACAGCGGCGAGGTGAACATCGTG 2160
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2095 ACCGAGCTGCAGGCCATCCAGTCGCAGGACAGCGGCGAGGTGAACATCGTG 2154

Qy   2161 ACCGACAGCCAGTACGCCCTGGGCATCAGCCTCAGGGCTGGAAGGGCAGC 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2155 ACCGACAGCCAGTACGCCCTGGGCATCAGCCTCAGGGCTGGAAGGGCAGC 2214

Qy   2221 CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAAGGTGTACCTGACCTGGGTG 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2215 CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAAGGTGTACCTGACCTGGGTG 2274

Qy   2281 CCCGCCCAAGGGCATCGGCCAGCTGCCCGAGAAGGACCAGATCGGC 2340
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2275 CCCGCCCAAGGGCATCGGCCAGCTGCCCGAGAAGGACCAGATCGGC 2334

Qy   2341 CGCAAGGTGCTGTTCCTGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGAC 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2335 CGCAAGGTGCTGTTCCTGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGAC 2394

Qy   2401 GACCTGTACGTGGGCAGCGGCCTGGATCGATTAAAAGCTTCCCGGGGCTAGCACC 2460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2395 GACCTGTACGTGGGCAGCGGCCTGGATCGATTAAAAGCTTCCCGGGGCTAGCACC 2454

Qy   2461 GGT 2463
          |||
Db   2455 GGT 2457

RESULT 7
ABL39961
ID   ABL39961 standard; DNA; 2457 BP.
XX
AC   ABL39961;
XX
DT   15-MAY-2002 (first entry)
XX

Qy   481 ATGAGCCTGCCCGGCAAGTGGAACGCCCAGATGATCGGGCGGCATCGGCGGCTTCATCAAG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCCGGCAAGTGGAACGCCCAGATGATCGGGCGGCATCGGCGGCTTCATCAAG 540

Qy   541 GTGCGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCCATCGGCGACCGTG 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGCGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCGACCGTG 600

Qy   601 CTGATCGGCGCCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCGCCCACCCCGTGAACATCATCGGCGCCGCAACATGCTGACCCAGCTGGGCTGC 660

Qy   661 ACCCTGAACTTCCCCATCAGCCCATCAGCCCGAGCGCGTGTGCCCGTGAAGCTGAAGCCCGGCATG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCCATCAGCCCATCAGCCCGAGCGCGTGTGCCCGTGAAGCTGAAGCCCGGCATG 720

Qy   721 GACGGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAATCAAGGCCCCTGACCGCC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAATCAAGGCCCCTGACCGCC 780

Qy   781 ATCTCCGAGGAGATGGAGAAGGAGGGCCAAGATCACCAAGATCGGCCCGAGAACCCCTAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTCCGAGGAGATGGAGAAGGAGGGCCAAGATCACCAAGATCGGCCCGAGAACCCCTAC 840

Qy   841 AACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAGCAGGACCACCAAGTGGCTGTGAC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAGCAGGACCACCAAGTGGCTGTGGAC 900

Qy   901 TTCCCGGAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGAGCATGCCCCCAC 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCCGGAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGAGCATGCCCCCAC 960

Qy   961 CCCGCCCGGCCTGGAAGAGCGTGACCGTGACCGTGCTGGACGTGGCGGACGCGCTACTTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCCCGGCCTGGAAGAGCGTGACCGTGACCGTGCTGGACGTGGCGGACGCGCTACTTC 1020

Qy   1021 AGCGTGCCCTGGAAGAGCACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCAAC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGCGTGCCCTGGAAGAGCACTTCCGCAAGTACACCACCGCCTTCACCATCCCCAGCATCAAC 1080

Qy   1081 AACGAGACCCCGGCCGATCCGCCGGCATCGCCGGCATCGTGCTGCCCGAGGGCTGAAGGGCAGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACGAGACCCCGGCCGATCCGCCGGCATCGCCGGCATCGTGCTGCCCGAGGGCTGAAGGGCAGC 1140

Qy   1141 CCCAGCATCTTCCAGAACGATGACCAAGATCGTCGACGACGCCCTTCCGCGGCGCCAACCCC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCCAGAACGATGACCAAGATCGTCGACGACGCCCTTCCGCGGCGCCAACCCC 1200

Qy   1201 GAGATCGTGATCTACCAGTACACGACGACCTGTAGGTGGGCAGCGACCTGGAGCAAGGGCATCGGC 1260
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCA------GGCCCCCGTGTACGTGGGCAGCGACCTGGAGCAAGGGCATCGGC 1254

Qy   1261 CAGCAGGACCCGCCCAAGATCGAGGAGACCTGCGCAAGACCTGCTGCGCCTTCACCACC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 CAGCAGGACCCGCCCAAGATCGAGGAGACCTGCGCAAGACCTGCTGCGCCTTCACCACC 1314

Qy   1321 CCCGACAAGAGCACCAGCAAGGAGCGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 CCCGACAAGAGCACCAGCAAGGAGCGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCC 1374

Qy   1381 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACGCTGGACCGTGAACGAC 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1375 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACGCTGGACCGTGAACGAC 1434

Qy   1441 ATCCAGAGAGCTGGTGGGCAAGCTGAACTGGGCGCGGCCAAGATCTACCCCGGCATCAAGGTG 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1435 ATCCAGAGAGCTGGTGGGCAAGCTGAACTGGGCGCGGCCAAGATCTACCCCGGCATCAAGGTG 1494

Qy   1501 CGCCAGCTGTGTCGCAAGCTGCTCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1495 CGCCAGCTGTGTCGCAAGCTGCTCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACC 1554

DE   Synthetic construct PR975YMWM SEQ ID NO:32.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO20020204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 10; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 U; 0 Other;

Query Match   97.8%; Score 2415.4; DB 6; Length 2457;
Best Local Similarity   99.3%; Pred. No. 2.3e-291;
Matches 2451; Conservative   0; Mismatches   6; Indels 12; Gaps   2;

Qy   1 GTCGACGCGCCACCATGGCCGAGGCCATGAGCGCCACCAGCGCGCCAACATCCTGATGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   1 GTCGACGCGCCACCATGGCCGAGGCCATGAGCGCCACCAGCGCGCCAACATCCTGATGCAG 60

Qy   61 CGGAGCAACTTCAAGGGCGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120
          ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db   61 CGGAGCAACTTCAAGGGCGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120

Qy   121 CACATGGCCCGCGACAACTGCCGCGGCGCCGCCGCCCCCGAAGAAAGGGCTGCTGAAGTGCGGGCAAGGAG 180
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CACATGGCCCGCGACAACTGCCGCGGCGCCGCCGCCCCCGAAGAAAGGGCTGCTGAAGTGCGGGCAAGGAG 180

Qy   181 GGCCACCAGCAGATGAAGGACTGCACCGAGCGGGTCCGGAGTTCCCGCGAGGCGCCCAACTTCTTCCGCGAGGACCTGGCC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 GGCCACCAGCAGATGAAGGACTGCACCGAGCGGGTCCGGAGTTCCCGCGAGGCGCCCAACTTCTTCCGCGAGGACCTGGCC 240

Qy   241 TTCCCCAGGGGCAAGGGCCCGCGAGTCGCGGCGACACCCCCGCAGCGAGGCGCCGGCCACC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TTCCCCAGGGGCAAGGGCCCGCGAGTTCCCGCGGCGACACCCCCGCAGCGAGGCGCCCACC 300

Qy   301 AGCGCGGAGCTGCAGGTCGCGGCGGCGGCGGAGGCCGGCGGCCGAGCGCCAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCGCCGGAGCTGCAGGTGCGGCGGTGCGGCGGACAACCCGGCCGGCCGAGCGCGGCCAG 360

Qy   361 GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGGCGCCCCCTGGTGAGCATCAAGGTG 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG 420

Qy   421 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAG 480

Qy   481 ATGAGCCTGCCGGCAAGTGAAGCCAAGATCGGCGGCATCGGCGGCTTCATCAAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGCCTGCCGGCAAGTGAAGCCAAGATCGGCGGCATCGGCGGCTTCATCAAG 540

Qy   541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCCACCGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCCACCGTG 600

Qy   601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGCTGGGCTGC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGCTGGGCTGC 660

Qy   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720

Qy   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC 780

Qy   781 ATCTGCGGAGGATGCGAGAGATGCGAGAGGCCAAGATCACCCAAGATCGGCCCCGAGAACCCCTAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTGCGGAGGATGCGAGAGATGCGAGAGGCCAAGATCACCCAAGATCGGCCCCGAGAACCCCTAC 840

Qy   841 AACACCCCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTCGATGGAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTCGATGGAC 900

Qy   901 TTCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTCGGCATCCCCAC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTCGGCATCCCCAC 960

Qy   961 CCCGCCGGCCTGAAGAAGAAGAAGAAGAAGCGGTGACCGTGCTGACAGTGGGCACGCGCTGCTACTTC 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCCGGCCTGAAGAAGAAGAAGAAGAAGCGGTGACCGTGCTGACAGTGGGCACGCGCTGCTACTTC 1020

Qy   1021 AGGGTGCCCCTGGACGAGGACTTCCGCAGAGCTACCGCCTTCACCATCCCCAGCATCAAC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AGGGTGCCCCTGGACGAGGACTTCCGCAGAGCTACCGCCTTCACCATCCCCAGCATCAAC 1080

Qy   1081 AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC 1140

Qy   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCTTCCCGGCCCGAACCCC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCTTCCCGGCCCGAACCCC 1200

Qy   1201 GAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGCAGTGCGGGAGATCGGC 1260
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1201 GAGATCGTGATCTACCA-----GGCCCCCGTGTACGTGGCAGTGCGGGAGATCGGC 1254

Qy   1261 CAGCACCGGCCAAGATCGAGGAGTGCGGCAAGCACCTGCTGCCTGGGGCTTCACCACC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 CAGCACCGGCCAAGATCGAGGAGTGCGGCAAGCACCTGCTGCCTGGGGCTTCACCACC 1314

Qy   1321 CCCGACAAGAAGCACCAGAAGGAGCCCCCTTCGTGGTGATGGCTACGAGCTGCACCCC 1380
          ||||||||||||||||||||||||||||||||||||||||||||| ----------
Db   1315 CCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCGTGGTGATGGCCCAT------CGGCTGCACCCC 1368

Qy   1381 GACAAGTGAAGCCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC 1440
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   1369 GACAAGTGAAGCCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC 1428

Qy   1441 ATCCAGAAGCTGGTGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db   1429 ATCCAGAAGCTGTGGTGGGCAAGCTGAACTGGGCGCGCCAAGGCCTGACCGACCATCGTGCCCCTGACC 1488

Qy   1501 CGCCAGCTGTGCAAGCTGCTCGCGGCGGCGGCGCCAAGGCCGCGCAAGGCCCGCGCGGGAGAGATCGGC 1560
           ||||||||||||||||||||||
Db   1489 CGCCAGCTGTGCAAGCTGCTCGCGGCGGCGGCGCCAAGGCCGCGCAAGGCCCGCGACCATCGTGCCCCTGACC 1548

Qy   1561 GAGGAGGGCCGAGCTGGGAGCTGGCGAGCTGGCCGAGATCCTGCGCGGAGACCGCGGCGAGATCCTGCGCGGAGATCCTGCGCACGGC 1620
           |||||||||||||||||||||||||||||||||||
Db   1549 GAGGAGGGCCGAGCTGGGAGCTGGCGAGCTGGCCGAGATCCTGCGCGGAGACCGCGGCGAGATCCTGCGCGGAGATCCTGCGCACGGC 1608

Qy   1621 GTGTACTACGACCCCAGCAAGGACCTGGTTGCCGGAGATCCAGAAGCAGGGCCACGACCAG 1680
           |||||||||||||||||||||||||||||||||||
Db   1609 GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGGAGATCCAGAAGCAGGGCCACGACCAG 1668

Qy   1681 TGGACTTACCAGATCTACCAGGAGAGCCCTTCAAGAACCTGAAGACCGCCAAGTACGCCAAG 1740
           |||||||||||||||||||||||||||||||||||
Db   1669 TGGACCTACCAGATCTACCAGGAGACCCCTTCAAGAACCTGAAGACCGCCAAGTACGCCAAG 1728

Qy   1741 ATGGCCACCGCCCACCACCAGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCC 1800
           |||||||||||||||||||||||||||||||||||
Db   1729 ATGGCCACCGCCCACCACCAGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCC 1788

Qy   1801 ATGGAGAGCATCGTGATCGTGGGCAAGACCCCCAAGTTCCGGCTGCCCATCCAGAAGGAG 1860
           |||||||||||||||||||||||||||||||||||
Db   1789 ATGGAGAGCATCGTGATCGTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1848

Qy   1861 ACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC 1920
           |||||||||||||||||||||||||||||||||||
Db   1849 ACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC 1908

Qy   1921 GTGAACACCCCCCCCCTGTGAAGCTGTGTGTACCAGCTGGAGAAGGAGCCCATCATCGGC 1980
           |||||||||||||||||||||||||||||||||||
Db   1909 GTGAACACCCCCCCCCTGGTGAAGCTGTGTGTACCAGCTGGAGAAGGAGCCCATCATCGGC 1968

Qy   1981 GCCGAGACCTTCTACGTGGACGAGGCCCGCCAACCGCGAGACCAAGATCGCCAAGGCCGC 2040
           |||||||||||||||||||||||||||||||||||
Db   1969 GCCGAGACCTTCTACGTGGACGAGGCGCGCCAACCGCGAGACCAAGATCGCGAAGGCCGC 2028

Qy   2041 TACGTGACCGACCGGGGCCGGCCGCAGAGATGTGAGCCTGACCGAGACCACCAACCAGAAG 2100
           |||||||||||||||||||||||||||||||||||
Db   2029 TACGTGACCGACCGGGGCCGGCCGCAGAGATCGTGAGCCTGACCGAGACCACCAACCAGAAG 2088

Qy   2101 ACCGAGCTGCAGGCCATCATCGGGCATGGCCCTGCCCTGCAGACAGCGGCAGCGGTGAACATCGTG 2160
           |||||||||||||||||||||||||||||||||||
Db   2089 ACCGAGCTGCAGGCCATCATCGAGCTGCCCTGCCCTGCAGACAGCGGCAGCGGTGAACATCGTG 2148

Qy   2161 ACCGACAGCAGCCAGTACGGCCCCGGCCATCCAGGCCCCGGCATCATCCAGGCCCCGGCAAGAGCGGAGCGGAG 2220
           |||||||||||||||||||||||||||||||||||
Db   2149 ACCGACAGCCAGTACGGCCCCGGCCATCCAGGCCCCGGCATCATCCAGGCCCCGGCAAGAGCGGAGCGGAG 2208

Qy   2221 CTGGTGCAACCAGATCATCGAGCAGCTGATCAAGAGAGGAGAGGTGTACCTGAGCTGGGTG 2280
           |||||||||||||||||||||||||||||||||||
Db   2209 CTGGTGCAACCAGATCATCGAGCAGCTGATCAAGAGAGGAGAGGTGTACCTGAGCTGGGTG 2268

Qy   2281 CCCGGCCCACAAGGGCATCGGCGGCAAGCAGCAGATCGACAAGCTGGTGAGCAAGGGCATC 2340
           |||||||||||||||||||||||||||||||||||
Db   2269 CCCGGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATC 2328

Qy   2341 CGCAAGGTGCTGTTCCTGGACGCGCATCGATGGCGCATCTGGATCGATGGCGGCTACCATGGAC 2400
           |||||||||||||||||||||||||||||||||||
Db   2329 CGCAAGGTGCTGTTCCTGGACGCGCATCGATGGCGCATCTCGATCGATGGCGGCTACCATGGAC 2388

Qy   2401 GACCTGTACGTGGGCAGCGGCGGCGGCGGCCCCTTCCCGGGCTAGCACC 2460
           |||||||||||||||||||||||||||||||||||
Db   2389 GACCTGTACGTGGGCAGCGGCGGCGGCGGCCCCTTCCCGGGCTAGCACC 2448

Qy   2461 GGTGAATTC 2469
           |||||||||
Db   2449 GGTGAATTC 2457
```

```
ID   ACA03546 standard; DNA; 2445 BP.
XX
AC   ACA03546;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Synthetic DNA ecoding immunogenic HIV peptide #29.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J,  Barnett SW,  Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 34; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 2445 BP; 562 A;  835 C;  751 G;  297 T;  0 U;  0 Other;

Query Match             97.3%;  Score 2401.8;  DB 7;  Length 2445;
Best Local Similarity   99.2%;  Pred. No. 1le-289;
Matches 2438;  Conservative    0;  Mismatches    7;  Indels    12;  Gaps    2;
```

```
Qy     7 GCCACCATGGCCGAGGCCATGAGCCAGCGCCAGCGCCAACATCCTGATGCAGCGCCAGC   66
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GCCACCATGGCCGAGGCCATGAGCCAGCGCCAGCGCCAACATCCTGATGCAGCGCCAGC   60

Qy    67 AACTTCAAGGGCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC  126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC  120

Qy   127 GCCCGCAACTGCGCGCCCCCCGCCAAGAAGGGCGCTGCTGGAAGTGCGGCAAGGAGGGGCCAC  186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCCGCAACTGCCGCGCCCCCGCCAAGAAGGGCGCTGCTGGAAGTGCGGCAAGGAGGGGCCAC  180

Qy   187 CAGATGAAGGACTGCACCGAGCGCCAGGGCCAACTTCTTCCGCGAGGACTTGGCCTTCCCC  246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGATGAAGGACTGCACCGAGCGCCAGGGCCAACTTCTTCCGCGAGGACTTGGCCTTCCCC  240
```

```
RESULT 8
ACA03546
```

```
Qy  1327  AAGAAGCACCAGAGGAGCCCCCTTCGTGTGATGGGCTACGAGCTGACCCCGACAAG 1386
          ||||||||||||||||||||||| |||| ||||| ||           | |||||||
Db  1315  AAGAAGCACCAGAGGAGCCCCCCTTCCTGTGCCCAT-----CGAGCTGCACCCCGACAAG 1368

Qy  1387  TGGACCGTGCAGCCCATCCAGTCGCCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG 1446
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1369  TGGACCGTGCAGCCCATCCAGTCGCCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG 1428

Qy  1447  AAGTTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTCGCGCCAG 1506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1429  AAGTTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAG 1488

Qy  1507  CTGTGCAAGCTGCTGCGCGGCCAAGGGCCTGACCGACATCGTGCCCCTGACCGAGGAG 1566
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  1489  CTGTGCAAGCTGCTGCGCGGCCAAGGGCCTGACCGACATCGTGCCCCTGACCGAGGAG 1548

Qy  1567  GCCGAGCTGGAGCTGGCCGAGAACCGGCAGATCCTGCGCGAGCCCGTGCACGGCCGTGTAC 1626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1549  GCCGAGCTGGAGCTGGCCGAGAACCGGCAGATCCTGCGCGAGCCCGTGCACGGCCGTGTAC 1608

Qy  1627  TACGACCCCCGCAAGGACCTGGTGGCCAAGGCCCTTCAAGAACCTGAAGACCGGCCACGACCAGTGGACC 1686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1609  TACGACCCCCGCAAGGACCTGGTGGCCAAGGCCCTTCAAGAACCTGAAGACCGGCCACGACCAGTGGACC 1668

Qy  1687  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAAGACCGGCCAAGTACGCCCAAGATGCGC 1746
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1669  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAAGACCGGCCAAGTACGCCCAAGATGCGC 1728

Qy  1747  ACCGCCCACACAACGACGTGAAGCAGTCTGACCGAGGCCGTGACCGAGAAGATCGCCATGGAG 1806
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1729  ACCGCCCACACAACCAGCGTGAAGCAGTCTGACCGAGGCCGTGACCGAGAAGATCGCCATGGAG 1788

Qy  1807  AGCATGTGATCTGGGCAAGACCCCCAAGTTCGCCTGGCCGATCCCGAGTGGGAGTTCGTGAAC 1866
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1789  AGCATGTGATCTGGGCAAGACCCCCAAGTTCGCCTGGCCGATCCCGAGTGGGAGTTCGTGAAC 1848

Qy  1867  GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGCCGCCGAG 1926
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1849  GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGCCGCCGAG 1908

Qy  1927  ACCCCCCCTGGTGTGGTGAAGCTGTGGTTACCAGCTGGAGAAGGAGCCCATCATCGGCCGCCGAGACCTG 1986
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1909  ACCCCCCCTGGTGTGGTGAAGCTGTGGTTACCAGCTGGAGAAGGAGCCCATCATCGGCCGCCGAGACCTG 1968

Qy  1987  ACTTCTACGTGGACGCGCCGCCAACCGGCGATCGGCGACACCAGAGAACCGGACCGTG 2046
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1969  ACTTCTACGTGGACGCGCCGCCAACCGGCGATCGGCGACACCAGAGAACCGGACCGTG 2028

Qy  2047  ACCGACCGGGCCGGCGCGAGAGATCGTGAGCCTGACCGACGCGGAGGTGAACATCGTGACCGAC 2106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2029  ACCGACCGGGCCGGCGCGAGAGATCGTGAGCCTGACCGACGCGGAGGTGAACATCGTGACCGAC 2088

Qy  2107  CTGCAGGAGTACGCCCTGGGCATCATCCAGGACAAGAGCCCGACAAGAGCGAGAGCGGAGACGAC 2166
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2089  CTGCAGGAGTACGCCCTGGGCATCATCCAGGACAAGAGCCCGACAAGAGCGAGAGCGGAGACGAC 2148

Qy  2167  AGCCAGTACGCCGATCCTGGGCCCAGGGCTTGGAAGGGCAGCCCCAGCCCGACAAGAGGCGAGAGCGGTG 2226
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2149  AGCCAGTACGCCGATCCTGGGCCCAGGGCTTGGAAGGGCAGCCCCAGCCCGACAAGAGGCGAGAGCGGTG 2208

Qy  2227  AACCAGATCATCGAGCAGCTGATCAAGAAGCAAGAGGAGAAGGTGTACCTGAGCTGGGTGCCGCC 2286
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2209  AACCAGATCATCGAGCAGCTGATCAAGAAGCAAGAGGAGAAGGTGTACCTGAGCTGGGTGCCGCC 2268

Qy  2287  CACAAGGGCATCGGCCGCAACGAGCAGAGGCATCGGCGCATCCGCAAG 2346
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  2269  CACAAGGGCATCGGCCGCAACGAGCAGAGCAGATGCAAGCGCTGAGCAAGGGCATCCGCAAG 2328

Qy  2347  GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTCGATCGTTGATCTCGTTACCAGTACGAGACCTG 2406
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2329  GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTCGATCGTTGATCTCGTTACCAGTACGAGACCTG 2388

Qy  2407  TACGTGGCAGCGGCGGCCGCGCCCTAGGATCGATTAAAAAGCTTCCCGGGCGATCGGGCTAGCACCGGT 2463
```

```
Qy  247   CAGGGCAAGGCCCGCGAGTTCCCCAGCGGAGCGAGCAACAGCCCCACCAGCCGC 306
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241   CAGGGCAAGGCCCGCGAGTTCCCCAGCGGAGCGAGCAACAGCCCCACCAGCCGC 300

Qy  307   GAGCTGCAGGTGCGCGGCGACAACCCCGAGCGCGAGGCCGGCCGGCCGAGCGGCCACC 366
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301   GAGCTGCAGGTGCGCGGCGACAACCCCGAGCGCGAGGCCGGCCGGCCGAGCGGCCACC 360

Qy  367   CTGAACTTCCCCCAGATCACCCTGGCCAGCGGCCGCCTGGTGAGCATCAAGGTGCGGCGGC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361   CTGAACTTCCCCCAGATCACCCTGGCCAGCGGCCGCCTGGTGAGCATCAAGGTGCGGCGGC 420

Qy  427   CAGATCAAGGAGGAGCCCTGCTGCTGCACCGTGCTGCTGAGGACACCGTGCTGGAGGAGATGAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421   CAGATCAAGGAGGAGCCCTGCTGCTGCACCGTGCTGCTGAGGACACCGTGCTGGAGGAGATGAGC 480

Qy  487   CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATGGCGGCCTTCATCAAGTGCTGCGC 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481   CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATGGCGGCCTTCATCAAGTGCTGCGC 540

Qy  547   CAGTACGACCAGATCTGATCGAGATCTCGGCCAAGAAGGCCATCGGCCACCGTGCTGCTGATC 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541   CAGTACGACCAGATCCTGATCGAGATCTCGGCCAAGAAGGCCATCGGCCACCGTGCTGCTGATC 600

Qy  607   GGCCCCACCCCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGTGGGCTGCACCCTG 666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601   GGCCCCACCCCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGTGGGCTGCACCCTG 660

Qy  667   AACTTCCCCATCAGCCCCATCAGCGCCGACGCGTGCCCGTGAAGCTGAAGCCGCGGACGGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661   AACTTCCCCATCAGCCCCATCAGCGCCGACGCGTGCCCGTGAAGCTGAAGCCGCATGGACGGC 720

Qy  727   CCCAAGGTGAAGCAGTGGCCCGAGGAGAAGATCAAGGCCCTGACCGGCCCATCTGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721   CCCAAGGTGAAGCAGTGGCCCGAGGAGAAGATCAAGGCCCTGACCGGCCCATCTGC 780

Qy  787   GAGGAGGATGGAAGGAGGGCCAAGATCACCAAGATCGGCCCCGAGACCCCTACACACC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781   GAGGAGGATGGAAGGAGGGCCAAGATCACCAAGATCGGCCCCGAGACCCCTACACACC 840

Qy  847   CCCGTGTTCGCCATCAAGAGAGGACAGCACAGCACCAAGGTGGCGCAGCTTCGCGC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841   CCCGTGTTCGCCATCAAGAGAGGACAGCACAGCACCAAGGTGGCGCAGCTTCGCGC 900

Qy  907   GAGCTGAACAAGCGCACCCCAGGACTTCTGGGAGGTGAGCTGGGCATCCCCCAGCCCCGCGC 966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901   GAGCTGAACAAGCGCACCCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAGCCCCGCGC 960

Qy  967   CCCCTGGACGAGGACTTCCGCAAGGCTTGACCGTGCTGGAGCGTGGGCGACGCTCAGCGTG 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961   GGCCTGGACGAGGACTTCCGCAAGGCTTGACCGTGCTGGAGCGTGGGCGACGCTACTTCAGCGTG 1020

Qy  1087  ACCCCGGCGATCCGCTACCAGTACAACGTGCTGCCCAAGGGCTGGAAGGGCAGCCCCAGC 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  ACCCCCGGCGATCCGCTACCAGTACAACGTGCTGCCCAAGGGCTGGAAGGGCAGCCCCAGC 1140

Qy  1147  ATCTTCCAGAGACAGCCATGACCAAGATCTGGAGCCCTTCCGGCGCCCAACCCGCAGATC 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  ATCTTCCAGAGACAGCAGCCATGACCAAGATCTGGAGCCCTTCCGGCGCCCAACCCGCAGATC 1200

Qy  1207  GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  GTGATCTACCA------GGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1254

Qy  1267  CGGCCCAAGATCGGAGGAGTCGAGGAGTTGCGCAAGCACCTGTGGGCTTCACCACCCCCGAC 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255  CGGCCCAAGATCGGAGGAGTCGAGGAGTTGCGCAAGCACCTGTGGGCTTCACCACCCCCGAC 1314
```

```
Db   2389 TACGTGGGGCAGCGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT 2445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 9
ADC13264
ID   ADC13264 standard; DNA; 2445 BP.
XX
AC   ADC13264;
XX
DT   18-DEC-2003  (first entry)
XX
DE   DNA of HIV construct p2Pol-opt-YMWM_C SEQ ID NO 43.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 40; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2445 BP; 562 A; 835 C; 751 G; 297 T; 0 U; 0 Other;

Query Match        97.3%;  Score 2401.8;  DB 9;  Length 2445;
Best Local Similarity  99.2%;  Pred. No. 1.1e-289;
Matches 2438;  Mismatches  0;  Conservative  0;  Indels 12;  Gaps  2;

Qy      7 GCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGGGCCAACATCCTGATGCAGCGGCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGGGCCAACATCCTGATGCAGCGGCAGC 60

Qy     67 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 120

Qy    127 GCCCGCAACTGCCGCCGCCCCCGCCGCCCCCCAAGAGAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC 186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCGCAACTGCCGCCGCCCCCGCCGCCCCCCAAGAGAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCAC 180
```

```
Qy    187 CAGATGAAGGACTGCACCGAGCGCGGCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAAGGACTGCACCGAGCGCGGCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC 240

Qy    247 CAGGGCCAAGGCCCGGCGAGTTCCCAGCAGCAGAACCGCGGCCAACAGCCCACCAGCCGC 306
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCCAAGGCCCGGCGAGTTCCCAGCAGCAGAACCGCGGCCAACAGCCCACCAGCCGC 300

Qy    307 GAGCTGCAGGTGCGCGGCACAACCCCGCGAGCGAGGGCCGGCGGCCGGCCGAGCGGCGGCCACC 366
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGCTGCAGGTGCGCGGCACAACCCCGCGAGCGAGGGCCGGCGGCCGGCCGAGCGGCGGCCACC 360

Qy    367 CTGAACTTCCCCAGATCACCCTGGCAGGGCCCCCTGGTGAGCATCAAGTGGGGCGGC 426
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCAGATCACCCTGTGCAGGGCCCCCTGGTGAGCATCAAGTGGGGCGGC 420

Qy    427 CAGATCAAGGAGGCCCTGCTGGACACCGGCGAGGACACCGTGCTGGAGGAGATGAGC 486
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGGCCCTGCTGGACACCGGCGAGGACACCGTGCTGGAGGAGATGAGC 480

Qy    487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGTGCGGC 546
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGTGCGGC 540

Qy    547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGAGGCCATCGGCACCGTGCTGATC 606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGAGGCCATCGGCACCGTGCTGATC 600

Qy    607 GGGCCCACCCCCCGCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGTGGGTGCACCCTG 665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGGCCCCACCCCCCGCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGTGGGGCTGCACCCTG 660

Qy    667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC 726
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC 720

Qy    727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGGCCATCTGC 786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGGCCATCTGC 780

Qy    787 GAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCACCAAGATCCGGCCCCGAGAACCCCTACAACACC 846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCACCAAGATCCGGCCCCGAGAACCCCTACAACACC 840

Qy    847 CCCGTGTTCGCCATCAAGAAGAAGAAGGACAGCACCAAGTGCGCAAGCTGGTGGACTTCCGC 906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCGTGTTCGCCATCAAGAAGAAGAAGGACAGCACCAAGTGCGCAAGCTGGTGGACTTCCGC 900

Qy    907 GAGCTGAACAAGGCGCACCCAGACTTCTGGGAGGTGCAGTGCAGCTGGCCATCCCCACCCCCGCC 956
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGCTGAACAAGGCGCACCCAGACTTCTGGGAGGTGCAGTGCAGCTGGCCATCCCCACCCCCGCC 960

Qy    967 GGCCTGAAGAAGAAGAAGAAGAGCGTGACCGTGGGGGCGACGCGCCTACTTCAGCGTG 1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCTGAAGAAGAAGAAGAAGAGCGTGACCGTGGGGGCGACGCGCCTACTTCAGCGTG 1020

Qy   1027 CCCCTGGACCGAGGACTTCCGCAGGATCACCGGCTTCACCATCCCCAGCATCAACAACGAG 1086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGAGGACTTCCGCAGGATCACCGGCTTCACCATCCCCAGCATCAACAACGAG 1080

Qy   1087 ACCCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCCAGCCCCAGC 1146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCCAGCCCCAGC 1140

Qy   1147 ATCTTCCAGACCAGCACGCATGACCAAGATCCTGGAGCCTTCCGCGCCGCAACCCGAGATC 1206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCTTCCAGACCAGCACGCATGACCAAGATCCTGGAGCCTTCCGCGCCGCAACCCGAGATC 1200

Qy   1207 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1266
          |||||||||||||||           |||||||||||||||||||||||||||||||||
Db   1201 GTGATCTACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1254

Qy   1267 CGGCCAAGATCGAGGAGCTGCGCCAAGCACTGCTGCGCTGGGGTTCACCACCCCCGAC 1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    2329 GTGCTGTTCCTGGACGGCGCGGCGGCCCTAGCATGCGGCATGCATCGATGCATGGCGGCATCGATCCTGATCTACCAGTACATGGACGACCTG 2388
Qy    2407 TACGTGGGCAGCGGCGGCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2463
Db    2389 TACGTGGGCAGCGGCGGCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2445

RESULT 10
ADC13230
ID   ADC13230 standard; DNA; 3930 BP.
XX
AC   ADC13230;
XX
DT   18-DEC-2003  (first entry)
XX
DE   DNA of HIV construct GagComplPolmut_C SEQ ID NO 9.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0318860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Lian Y,  Engelbrecht S,  Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 6; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 890 A; 1365 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match             97.0%;  Score 2394.8;  DB 9;  Length 3930;
Best Local Similarity   99.2%;  Pred. No. 7.8e-289;
Matches 2431; Conservative    0; Mismatches    7; Indels   12; Gaps    2;

Qy     14 TCGGCGGAGGCCATGAGCCATGAGCGGCCACCAGCGCCAAACATCCTGATGCAGCGCAGCAACTTCA 73
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 TCGGCGGAGGCCATGAGCCATGAGCGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA 1546

Qy     74 AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCGCA 133
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCGCA 1606

Qy    134 ACTGCCGGCCCGCCCCCCCGGCAAGGAGGGGGCTGCTGGAAGTGCGGCGGCACCAGATGA 193
```

```
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 CGGGCCAAGATCGAGGAGAGGAGCTGCCAAGCACCTGCTGCGTGGGCCTTGCTGTGGAATGGGCTACGAGCTGCACCGGCCACCCCCGACAAG 1314
Qy   1327 AAGAAGCACCAGAAGACCACCAGAAGAGGAGCTGCCAAGCACCTGCTGCGTGTGGAATGGGCTACGAGCTGCACCGGCCACCCCCGACAAG 1386
Db   1315 AAGAAGCACCAGAAGACCACCAGAAGAGGAGCTGCCAAGCACCTGCTGCGTGTGGAATGGGCTACGAGCTGCACCGG-----CGAGCTGCACCCCGACAAG 1368
Qy   1387 TGGACCGTGCAGCCATCCAGCTGCCCGGAAGGAGAGCTGGACCGTGAACGACATCCAG 1446
Db   1369 TGGACCGTGCAGCCATCCAGCTGCCCGGAAGGAGAGCTGGACCGTGAACGACATCCAG 1428
Qy   1447 AAGCTGGTGGGGCAAGCTGAACTGGGCCAGATTCACCCGGCATCAAGGTGCGCCAG 1506
Db   1429 AAGCTGGTGGGGCAAGCTGAACTGGGCCAGATTCACCCGGCATCAAGGTGCGCCAG 1488
Qy   1507 CTGTGCAAGCTGCTGCGCGGCGGCGCCAAGGCCCTGACGACATCGTGCCCCTGACCGAGGAG 1566
Db   1489 CTGTGCAAGCTGCTGCGCGGCGGCGCCAAGGCCCTGA--CGACATCGTGCCCCTGACCGAGGAG 1548
Qy   1567 GCCGAGCTGGAGAGCTGGCCGAGAACCGGCGAGATCTGCGGCGAGCCCGTGCACGGCGTGTAC 1626
Db   1549 GCCGAGCTGGAGAGCTGGCCGAGAACCGGCGAGATCTGCGGCGAGCCCGTGCACGGCGTGTAC 1608
Qy   1627 TACGACCCCAGCAAGAACGACCTGTGGTGGCGCAAGGAGGCCACGACCAGTCAGTGACC 1686
Db   1609 TACGACCCCAGCAAGAACGACCTGTGGTGGCGCAAGGAGGCCACGACCAGTCAGTGACC 1668
Qy   1687 TACCAGATCTACCAGGAGCCTTCAAGGAACCTGAAGACCGGCCAAGTACGCCGCCAAGATCGC 1746
Db   1669 TACCAGATCTACCAGGAGCCCTTCAAGGAACCTGAAGACCGGCCAAGTACGCCGCCAAGATCGC 1728
Qy   1747 ACCGCCACCAACGACGTGAAGACCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1806
Db   1729 ACCGCCCA-CACCAACGACGTGAAGACCAGCTGACCGAGGCCGTGCAGAAGATCGCCATCGAG 1788
Qy   1807 AGCATCGTGATCTGGGGCAAGACCACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAGACCTGG 1866
Db   1789 AGCATCGTGATCTGGGGCAAGACCACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAGACCTGG 1848
Qy   1867 GAGACCTCGTGACCGACTACTGCGCAGGCCACCTGGACTGGGAGTTCGTGAAC 1926
Db   1849 GAGACCTCGTGACCGACTACTGCGCAGGCCACCTGGACTGGGAGTTCGTGAAC 1908
Qy   1927 ACCCCCCCCTGGTGAAGCTGTGTGTACCAGCTGAGAAGGAGAGCCCATCATCGGCGCCCGAG 1986
Db   1909 ACCCCCCCCTGGTGAAGCTGTGTGTACCAGCTGAGAAGGAGAGCCCATCATCGGCGCCCGAG 1968
Qy   1987 ACCTTCTACGTGGACGGCGGCCGCCAACCGCGAGACCAAGATCGCCAAGGCCGCTACGTG 2046
Db   1969 ACCTTCTACGTGGACGGCGGCCGCCAACCGCGAGACCAAGATCGCCAAGGCCGCTACGTG 2028
Qy   2047 ACCGACCGGGGCGGCCAGAAGATCGTGACCAAGACCGAGCTGAGCCTGACCGACACCAGAAGACCGAG 2106
Db   2029 ACCGACCGGGGCCGCCAGAAGATCGTGACCAAGACCGAGCCTGACCGACACCAACCAGAAGACCGAG 2088
Qy   2107 CTGCAGGCCCATCCAGCTGGCCCCTGCAGGACAGCAGCGCAGCGGCGAGCTGAACATCGTGACCGAC 2166
Db   2089 CTGCAGGCCCATCCAGCTGGCCCCTGCAGGACAGCAGCGGCGAGCTGAACATCGTGACCGAC 2148
Qy   2167 AGCCAGTACGCCCTGGGCATCATCAGGCCCAGCCCAGACAAGAGCCAGAGCGAGCTGGTG 2226
Db   2149 AGCCAGTACGCCCTGGGCATCATCAGGCCCAGCCCAGACAAGAGCCAGAGCGAGCTGGTG 2208
Qy   2227 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2286
Db   2209 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCCCC 2268
Qy   2287 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2346
Db   2269 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2328
Qy   2347 GTGCTGTTCCTGGACGGCGCGGCATCGATGGCGGCATCGATCCTGATCTACCAGTACATGGACGACCTG 2406
```

```
Db   2687 ACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGTCGGCAGCACCGCGCCCA 2740
          ||||      ||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1274 AGATGGAGGAGCTGCGCAAGCACCTGCGTGGGGCTTCACCCACCCCGACAAGGAGC 1333
          ||||||||||||||||||||||||||||||||||||||||| ||||||||| |||||
Db   2741 AGATGGAGGAGCTGCGCAAGCACCTGCGTGGGGCTTCACCCACCCCGACACAGGAGC 2800

Qy   1334 ACCAGAAGGAGCCCCCTTCCTGTGGGATGGGCTACGAGCTGCACCCGACAAGTGGACCG 1393
          |||||||||||||||||||||||||||||      ||||||||||||||||||||||||
Db   2801 ACCAGAAGGAGCCCCCTTCCTGCCAT------CGAGCTGCACCCGACAAGTGGACCG 2854

Qy   1394 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCGTGGACCGTGACGAGCTGAACGACTGG 1453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2855 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCGTGGACCGTGACGAGCATCAGAAGCTGG 2914

Qy   1454 TGGGCAAGTGAACTGGGCGCAGATCTACCCGGCATCAAGGTGGCGGCCAGTGTGCA 1513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2915 TGGGCAAGTGAACTGGGCGCAGATCTACCCGGCATCAAGGTGGCGGCCAGTGTGCA 2974

Qy   1514 AGTGCTGCGCGGGCGCCAAGGCCCCTGACCGACATCGTGCCCCTGACCGAGGAGCCGAGC 1573
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2975 AGTGCTGCGCGGGCGCCAAGGCCCCTGACCGACATCGTGCCCCTGACCGAGGAGCCGAGC 3034

Qy   1574 TGGAGCTGGCCGAGAACCGGCAGATCCTGCGCAGCCCGTGCACGGCGGCGTGTACTACGACC 1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3035 TGGAGCTGGCCGAGAACCGGCAGATCCTGCGCAGCCCGTGCACGGCGGCGTGTACTACGACC 3094

Qy   1634 CCAGCAAGGACCTGGTGCGGCCAAGAGGCAGGGCCACAGTGGACTTACCAGA 1693
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3095 CCAGCAAGGACCTGGTGCGGCCAAGAGGCAGGGCCACAGTGGACTTACCAGA 3154

Qy   1694 TCTACCAGGAGACCCCTTCAAGAACCTGAAGACCCGGCCAAGTACGGCCACCCGCCC 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3155 TCTACCAGGAGACCCCTTCAAGAACCTGAAGACCCGGCAAGTACGGCCACCCGCCC 3214

Qy   1754 ACACCAACGACGTGAAGCAGCAGCGTGACCGAGGATCGCCATGACGAGGAGCATCG 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3215 ACACCAACGACGTGAAGCAGCAGCGTGACCGAGGATCGCCATGACGAGGAGCATCG 3274

Qy   1814 TGATCTGGGGCAAGACACCCCCAAGTTCGGCTGACCCATCTGCGAGGAGACTGGGAGACCT 1873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3275 TGATCTGGGGCAAGACACCCCCAAGTTCGGCTGACCCATCTGCGAGGAGACTGGGAGACCT 3334

Qy   1874 GGTGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGGAGTTCGTGAACACCCCCC 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3335 GGTGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGGAGTTCGTGAACACCCCCC 3394

Qy   1934 CCCTGGTGAAGCTGTGTGTACCAGCTGGAGAAGGGACCCATCATCCGCGGCGAGACCTTCT 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3395 CCCTGGTGAAGCTGTGTGTGGTACCAGCTGGAGAAGGGACCCATCATCCGCGGCGAGACCTTCT 3454

Qy   1994 ACGTGGACGGCGCCGCCCAACCGGAGACGAGATCGGCAAGGGCGGCTACGTGACCGACC 2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3455 ACGTGGACGGCGCCGCCCAACCGGAGACGAGATCGGCAAGGGCGGCTACGTGACCGACC 3514

Qy   2054 GGGGCCGGCCAGCAGATCTGAGCGGAGCAGATCGACCAGAAGACGAGCTGACGAGTGACGAGG 2113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3515 GGGGCCGGCCAGCAGATCTGAGCGGAGCAGATCGACCAGAAGACGAGCTGACGAGTGACGAGG 3574

Qy   2114 CCATCCAGCTGGCCCTGCAGGACACCGGGCAGCATCAACATCTGCAGAGCAGCCAGT 2173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3575 CCATCCAGCTGGCCCTGCAGGACACGCGGCAGCATCAACATCTGCAGACACAGCCAGT 3634

Qy   2174 ACGCCCTGGGCATCATCCAGGGCGACTCAGGGCTGGAAGGCGAGTGGAACCAGA 2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3635 ACGCCCTGGGCATCATCCAGGGCGACTCAGGGCTGGAAGGCGAGTGGAACCAGA 3694

Qy   2234 TCATGCAGCAGCATGACCATGACCAAGATCGTTGATCT 2293
          ||||||||||||||||||||||||||||||||||||||
Db   3695 TCATGCAGCAGCATGACCATGACCAAGATCGTTGATCT 3754

Qy   2294 GCATCGGCGGCAAGCAGCAGATCGACAAGGGCATCCCGCCA 2353
          ||||||||||||||||||||||||||||||||||||||||||
Db   3755 GCATCGGCGGCAAGCAGCAGATCGACAAGGGCATCCCGCCA 3814
```

```
Db   1607 ACTGCGCCGCGCCCCCGGAAGAGGGCTGCTGGAAGTCGGCAAGGAGGGGCCACCCAGGGCA 1666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    194 AGGACTGCACCGAGCGGCGCCAGGCCAACTTCTTCCGCGGACGGGACTGCTGCCCCAGGGCA 253

Db   1667 AGGACTGCACCGAGCGGCGCCAGGCCAACTTCTTCCGCGGACGGGACTGCTGCCCCAGGGCA 1726
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    254 AGGCCCGGCAGTTCCCCAGCGAGCAGAACCGGCCAACAGCCCCACCAGCCGCGAGCTGC 313

Db   1727 AGGCCCGGCAGTTCCCCAGCGAGCAGAACCGGCCAACAGCCCCACCAGCCGCGAGCTGC 1786
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    314 AGTGCGCGGCGACAACAACCGGCCGCGAGGCCGGCCGCGAGGGGCCAAGGGCACCTGAACT 373

Db   1787 AGTGCGCGGCGACAACAACCGGCCGCGAGGCCGGCCGCGAGGGGCCAAGGGCACCTGAACT 1846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    374 TCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTGGCGGCCAGATCA 433

Db   1847 TCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTGGGCGGCCAGATCA 1906
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    434 AGGAGGCCCTGCTGACACCGGCCAGCACACCGGCGCCGCGAGGATAGCGGCCTCCCG 493

Db   1907 AGGAGGCCCTGCTGACACCGGCCAGCACACCGGCGCCGACGAGCGAGCGATGAGCCTCCCCA 1966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    494 GCAAGTGGAAGCCAAGATGATCGGGCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG 553

Db   1967 GCAAGTGGAAGCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG 2026
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    554 ACCAGATCCTGATCGAGAGGCAAGGCCATCGGCATCGGCACCGTGCTGATCGGCGCCCCA 613

Db   2027 ACCAGATCCTGATCGAGAGGCAAGGCCATCGGCATCGGCACCGTGCTGATCGGCGCCCCA 2086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    614 CCCCGTGAACATCATCGGCGCCAACATGCTGACCACCAGTCGGCTGGCTGACCCTGAACTTCC 673

Db   2087 CCCCGTGAACATCATCGGCGCCAACATGCTGACCACCAGTCGGCTGGGCTGGCTGACCCTGAACTTCC 2146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    674 CCATCAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGAAGCCCGGCCCCAAGG 733

Db   2147 CCATCAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGAAGCCCGGCCCCAAGG 2206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    734 TGAAGCAGTGGCCCTGACCGAGGAGAAGCTCGGGGACTTCTTGGGAGTCCTGCGAGGAGA 793

Db   2207 TGAAGCAGTGGCCCTGACCGAGGAGAAGCTCTGGGAGATCGGGGACTTCTTGGGAGTCCTGCGAGGAGA 2266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    794 TGGAGAGGAGGGCAAGATCACCACCAAGATCGGCCCCCGAGAACCCCGGCAAACCCCCCGTGT 853

Db   2267 TGGAGAGGAGGGCAAGATCACCACCAAGATCACCAACCGGCCCCCGAGAACCCCGGCAAACACCCCCGTGT 2326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    854 TCGCCATCAAGGAAGGACGACCACCAAGCTGGTGTGACTTCCGGACTTCCGCGAGCTGA 913

Db   2327 TCGCCATCAAGGAAGGACAGCACCACCAAGCTGGTGTGACTTGGCACTTCCGCGAGCTGA 2386
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    914 ACAAGGCGCACCCCAGGACTTCTGGGAGGTGCACTCGACTGGAGTGTCGAGCGGCCGGTGA 973

Db   2387 ACAAGGCGCACCCCAGGACTTCTGGGAGGTGCACTCGACTGGAGTGCGCCAACCCCACCCCACCCCGGCCGGTGA 2446
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    974 AGAAGAAGAAGAGGCTTCCGCCAAGATCGCCAAGATCACCACCAAGATCCCCAGCATCGTTACTTTCC 1033

Db   2447 AGAAGAAGAAGAGGCTTCCCCAAGATCGCCAAGATCACCACCAAGATCCCCAGCATCGTTACTTTCC 2506
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1034 ACGAGGACTTCCGCCAAGCATCACCGTGGAGCGGCTGGAAAGGGCAGCTCCCAGCATCTGGG 1093

Db   2507 ACGAGGACTTCCGCCAAGCATCACCGTGGAGCGGCTGGAAGGGCAGCTCCCAGCATCTGGG 2566
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1094 GCATCCGTACCAGTACAACGTGCTGCCCGAGCGGCCTGGAAGGGCAGCCCCAGCATCTTCC 1153

Db   2567 GCATCCGTACCAGTACAACGTGCTGCCCGAGCGGCCTGGAAGGGCAGCCCCAGCATCTTCC 2626
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1154 AGAGCAGCAGCATGACCAAGATCACCTGGAGCTCAGGAAGAGGAAGAGGACTGTCGATCT 1213

Db   2627 AGAGCAGCAGCATGACCAAGATCACCTGGAGCTTCAGGAAGAGGAAGAGGACTGTCGATCT 2686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1214 ACCAGTACATGGAGCGACCTGTACGTGGGCAGCGACCTGGAGTCGGCAGCACCGCGCCCA 1273
```

RESULT 11
ADC13231
ID   ADC13231 standard; DNA; 3930 BP.
XX
AC   ADC13231;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagComplPolmutAlt_C SEQ ID NO 10.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-016860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 7; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramuscosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 889 A; 1365 C; 1214 G; 462 T; 0 U; 0 Other;

Query Match   96.9%;   Score 2393.2; DB 9;   Length 3930;
Best Local Similarity 99.2%;   Pred. No. 1.2e-288;
Matches 2430; Conservative 0; Mismatches 8;   Indels 12; Gaps 2;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCGGCCACCAGCGCCAACATCCTGATGCAGCGGCAGCAACTTCA   73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 TCGCCGAGGCCATGAGCCAGGCCACCGGCCACCAGCGCCAACATCCTGATGCAGCGCGCAGCAACTTCA   1546

Qy   74 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCCGCA   133
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627 AGAGCAGCAGCAATGACCAAGATCCTGAGACCCTTCCGGCCCCCAACCCCAACCCCGAGATCGTGATCT   2686

Qy   1547 AGGGCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCCGCA   1606
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   134 ACTGCGGCGGCCCCGCCCCCAAGAAGGGCTGTGGAAGTGCGGCAAGAGGAGGGCCACCAGATGA   193

Db   1607 ACTGCGGCGGCCCCGCCCCCAAGAAGGGCTGTGGAAGTGCGGCAAGAGGAGGGCCACCAGATGA   1666

Qy   194 AGGACTGCACCGAGCGGCCAGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCA   253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667 AGGACTGCACCGAGCGGCCAGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGGCA   1726

Qy   254 AGGGCCCGGAGTTCCCCAGCGAGGACAGCGCCAACAGCCCAACCAGCCCGAGCTGC   313
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727 AGGGCCGGAGTTCCCCAGCGAGGACAGCGCCAACAGCCCCACCAGCCGCGGCCGAGCTGC   1786

Qy   314 AGGTGCGGCGGCCGACAACCCCCGCCGGAGCGAGCGCCAGCGCCAGCGGCCACCCTGAACT   373
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787 AGGTGCGGCGGCCGACAACCCCCGCCGGAGCGAGCGCCAGCGCCAGCGGCCACCCTGAACT   1846

Qy   374 TCCCCCAGATCACCCTGTGGGAGCGGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA   433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847 TCCCCCAGATCACCCTGTGGGAGCGGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA   1906

Qy   434 AGGAGGCCCTGCTGGACACCGGCGCGCCGACGACACCGGTGGGAGGAGATGAGCCTGCCCG   493
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907 AGGAGGCCCTGCTGGACACCGGCGCGCCGACGACACCGGTGGGAGGAGATGAGCCTGCCCG   1966

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAGGTGCGCCAGTACG   553
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAGGTGCGCCAGTACG   2026

Qy   554 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATGGCCCGCGTGCTGATCGGCCCCA   613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATGGCCCGCGTGCTGATCGGCCCCA   2086

Qy   614 CCCCGTGACAACATCATCGGCCGCAACATGCTGACCCAGTGGGGCTGCACCCTGAACTTCC   673
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2087 CCCCGTGACATCATCGGCCGCAACATGCTGACCCAGTGGGGCTGCACCCTGAACTTCC   2146

Qy   674 CCATCAGCCCCATCGAGACCCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG   733
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2147 CCATCAGCCCCATCGAGACCCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG   2206

Qy   734 TGAAGCAGTGCCCCTGACCGAGAAGATCAAGGCCTGACCGCCATCGCCACCGAGGAGA   793
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2207 TGAAGCAGTGCCCCTGACCGAGAAGATCAAGGCCCTGACCGCCATCGCCACCGAGGAGA   2266

Qy   794 TGGGAGGAGGGCAAGATCACCCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGT   853
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2267 TGGGAGGAGGGCAAGATCACCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGT   2326

Qy   854 TCGCCATCAAGAAGAAGACAGCAGCAGTGGCGCCAAGTGGTGGACTTCCGGCGGCTGA   913
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327 TCGCCATCAAGAAGAAGACAGCAGCAGTGGCCCAAGTGGTGGACTTCCGGGCGGCTGA   2386

Qy   914 ACAAGCGCACCCAGACTTCTGGAGGTGCAGCTGGGCATCCCGCCGGCCTGA   973
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387 ACAAGCGCACCCAGACTTCTGGAGGTGCAGCTGGGCATCCCCACCCGCCGCCCTGA   2446

Qy   974 AGAAGAAGAGGAGCGGTGACCGTGACCGTGACCGTGGGGCGACGGCTACTTCAGCGTGCCCCTGG   1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2447 AGAAGAAGAGGAGCGGTGACCGTGACCGTGACCGTGGGGCGACGGCTACTTCAGCGTGCCCCTGG   2506

Qy   1034 ACGAGGACTTCCGCAAGTACACCGCCTTCAACATCCCAGCATCCCAGCATCAACAACGAGACCCCG   1093
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2507 ACGAGGACTTCCGCAAGTACACCGCCTTCAACATCCCAGCATCCCAGCATCAACAACGAGACCCCG   2566

Qy   1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCATCTTCCC   1153
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2567 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCATCTTCCC   2626

Qy   1154 AGAGCAGCATGACCAAGATCCTGAGACCCTTCCGGCCCCCAACCCCGAGATCGTGATCT   1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627 AGAGCAGCATGACCAAGATCCTGAGACCCTTCCGGCCCCCAACCCCAACCCCGAGATCGTGATCT   2686

Qy   2354 TCCTGGACGGCATCGATGCGGCGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG   2413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3815 TCCTGGACGGCATCGATGCGGCGCATCGTGATCTACCAGTACACTGGACGACCTGTACGTGG   3874

Qy   2414 GCAGCGGCGGCGGCCCTAGGATCGATTAATAAAAGCTTCCCGGGGGCTAGCACCGGT   2463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3875 GCAGCGGCGGCGGCCCTAGGATCGATTAATAAAAGCTTCCCGGGGGCTAGCACCGGT   3924

```
Qy  2294 GCATCGGCGGCAACGAGCAGCAGATCGACAAGCTGGTGGAGCAAGGGCATCCGCAAGGTGCTGT 2353
         ||||||||||||||||||||||||  ||||||||||||| |||||||||||||||||||||||
Db  3755 GCATCGGCGGCAACGAGCAGCAGATGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGT 3814

Qy  2354 TCCTGGACGGCATCGATGGCGGCATCGGCATCGATGGCGGCATCGATGGCGGCATGGACGACCTGTACGTGG 2413
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3815 TCCTGGACGGCATCGATGGCGGCATCGGCATCGATGGCGGCATCGATGGCGGCATGGACGACCTGTACGTGG 3874

Qy  2414 GCAGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2463
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3875 GCAGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 3924
```

RESULT 12
ADC13232
ID   ADC13232 standard; DNA; 3930 BP.
XX
AC   ADC13232;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagComplPolmutina_C SEQ ID NO 11.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO200300462O-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 8; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intranasally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenous-y. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 889 A; 1366 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match   96.9%; Score 2393.2; DB 9; Length 3930;
Best Local Similarity   99.2%; Pred. No. 1.2e-288;
Matches 2430; Conservative 0; Mismatches 8; Indels 12; Gaps 2;

Qy   14 TGGCCGAGGCCATGCAGCCAGGGCCAACATCCTGATGCAGCGGCAGCAACTTCA 73
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1487 TCGCCGAGGCCATGAGCCAGGGCCAACATCCTGATGCAGCGGCAGCAACTTCA 1546
```

```
Qy  1214 ACCAGTACATGGACGACCTGTACGTGGGCGAGGAGAGCCTGAGCGAGCCGCGCCA 1273
         ||||        ||||||||||||||||||||||||||||||||||||||||||||
Db  2687 ACCA------GGCCCCCGTGTACGTGGGCGAGGAGAGCCTGAGCGAGCCGCGCCA 2740

Qy  1274 AGATCGAGGAGCTGGCCAAGCACCTGCTGCGGGCTTCACCACCCCGACAAGAGAGC 1333
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2741 AGATCGAGGAGCTGGCCAAGCACCTGCTGCGGGCTTCACCACCCCGACACAAGAGAGC 2800

Qy  1334 ACCAGAAGGAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGACAAGTGGACCG 1393
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2801 ACCAGAAGGAGGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCGACAAGTGGACCG 2854

Qy  1394 TGCAGCCCATCGAGCTGCCCGAGAAGGAGGAGAGCGACCTGAACGACAATCCAGAAGCTGG 1453
         ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2855 TGCAGCCCATCGAGCTGCCCGAGAAGGAGGAGAGCGACCTGAACGACAATCCAGAAGCTGG 2914

Qy  1454 TGGGCCAAGCTGAACTCGGGCGGCGCCAGCATCTTACCCCGGCATCAAGGTGCGGCCAGCTGTGCA 1513
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2915 TGGGCCAAGCTGAACTCGGGCGGCGCCAGCATCAACAAGGTGCGGCCAGCTGTGCA 2974

Qy  1514 AGCTGCTCGCGGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCGAGC 1573
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2975 AGCTGCTCGCGGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGGCCGAGC 3034

Qy  1574 TGGAGCTGGCCGAGAACCGGCGAGATCTGCCGGAGCCCGTGCACGGCCGTGTACTACGACC 1633
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3035 TGGAGCTGGCCGAGAACCGGCGAGATCTGCCGGAGCCCGTGCACGGCCGTGTACTACGACC 3094

Qy  1634 CCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGA 1693
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3095 CCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGA 3154

Qy  1694 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCCAAGATGCGCCAAGATGCGCCACCGCC 1753
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3155 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCCAAGTGCGACCGGCCAAGTGCGACCGCCC 3214

Qy  1754 ACACCAACGACGTGAAGCAGCTGACTGACCGAGAAGATCGCCATCGGCGCGCCGAGCATCG 1813
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3215 ACACCAACGACGTGAAGCAGCTGACTGACCGAGAAGATCGCCATCGGCGCGCCGAGCATCG 3274

Qy  1814 TGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCAGTCCCCAGAAGGAGGAGACTGGGAGACCT 1873
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3275 TGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCAGTCCCCAGAAGGAGGAGACTGGGAGACCT 3334

Qy  1874 GGTGGACCGACTACTGCAGGCCACCTGCAGGGAGTTCGTGAAACACCCCCC 1933
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3335 GGTGGACCGACTACTGCAGGCCACCTGCAGGGAGTTCGTGAAACACCCCCC 3394

Qy  1934 CCCTGGTGAAGCTGTGTGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 1993
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3395 CCCTGGTGAAGCTGTGTGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 3454

Qy  1994 ACGTGGACGGCGGCGCCAACCGCGAGACCGAGACCAAGATCGGCAAGGCCGGCTACGGTGACCACC 2053
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3455 ACGTGGACGGCGGCGCCAACCGCGAGACCGAGACCAAGATCGGCAAGGCCGGCTACGGTGACCGACC 3514

Qy  2054 GGGGCCGGCCAGAAGATCGTGAGCTGACCGGCCTGACCGGCCAACCACCAGAGACCGAGCTGCAGG 2113
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3515 GGGGCCGGCCAGAAGATCGTGAGCTGACCGGCCTGACCGGCCAACCACCAGAGACCGAGCTGCAGG 3574

Qy  2114 CCATTCCAGCTGCCCTGCCCCTGCCCGAGAACAGCAGCCGCAGCGAGTGAACATCGTGACCGACCAGT 2173
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3575 CCATTCCAGCTGCCCTGCCCCTGCCCGAGAACAGCAGCCGCAGCGAGTGAACATCGTGACCGACCAGT 3634

Qy  2174 ACGCCCTGGGCATCATCCAGGCCCAGCCGCAGCAGCTCAAGAGCGAGAGCGAGCGGCAGTCGTGACCAGA 2233
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3635 ACGCCCTGGGCATCATCCAGGCCCAGCCGCAGCAGCTCAAGAGCGAGAGCGAGCGGCAGCTGTGAACCAGA 3694

Qy  2234 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCCGCCACAAGG 2293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3695 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCCGCCACAAGG 3754
```

Nucleotide sequence alignment (Query "Qy" vs Database "Db").

Right-hand column (positions 1154–3334 / 1213–3694):

```
Qy  1154  AGAGCAGCATGACCAAGATCTGCTGGAGCCCTTCCGCGGCCCGCAACCCGAGATCGTGATCT  1213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2627  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCCGCAACCCGAGATCGTGATCT   2686

Qy  1214  ACCAGTACATGGACGACCTGTACTGGGCAGCGACCTGGCAGCGACCGCGCCA           1273
          |||||||||||         |||||||||||||||||||||||||||||||
Db  2687  ACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCACGCCA          2740

Qy  1274  AGATCGAGGAGTCGCAAGCACCTGCTGCGGCTTCACCACCCCGACAAGAGC           1333
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2741  AGATCGAGGAGTCGCGCAAGCACCTGCGCGGCTTCACCACCCCGACAAGAGC          2800

Qy  1334  ACCAGAAGGAGCCCCCTTCCTGTGGATGGGTACGAGCTGCACGAGCTGGACCG          1393
          ||||||||||||||||||||||||       
Db  2801  ACCAGAAGGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCGACAAGTGGACCG      2854

Qy  1394  TGCAGCGCCATCGAGCTGCCCAGAAGGAGAGACTGGACCGTGAACGACATCCAGAAGCTGG    1453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2855  TGCAGCGCCATCGAGCTGCCCGAGAAGGAGAGACTGGACCGTGAACGACATCCAGAAGCTGG   2914

Qy  1454  TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA    1513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2915  TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA    2974

Qy  1514  AGCTGCTGCGCGGCGACACCGGCCAAGGCCCTGACCGTGCCCCTGACCGTGCCCGAGAGCC   1573
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2975  AGCTGCTGCGCGGCGACACCGGCCAAGGCCCTGACCGTGCCCCTGACCGTGCCCGAGAGC   3034

Qy  1574  TGGAGTCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACC   1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3035  TGGAGTCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACC   3094

Qy  1634  CCAGCAAGGACCTGGTGCGCCGAGATCCAGAGAACCGGCCACGAGCAGGCCACGCGCCC     1693
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3095  CCAGCAAGGACCTGGTGCGCCGAGATCCAGAGAACCGGCCACGAGCAGGCCACCGCGCCCC   3154

Qy  1694  TCTACCAGGAGAGCCCTTCAAGAACCTGAAGAACCTGAAGCAGAGTGCGCAAGATCG       1753
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3155  TCTACCAGGAGAGCCCTTCAAGAACCTGAAGAACCTGAAGCAGAGTACGGCAAGATCG      3214

Qy  1754  ACACCAACGACGTGAAGACACCCCAAGTTCCGCCTGCCCATCCAGAGGAGACCTGGGAGACCT   1813
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3215  ACACCAACGACGTGAAGACACCCCAAGTTCCGCCTGCCCATCCAGAGGAGACCTGGGAGACCT   3274

Qy  1814  TGATCTGGGCAAGACACCCCAAGTCCCGAGTCGTTGTGAACACCCCCC              1873
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3275  TGATCTGGGCAAGACACCCCAAGTCCCGAGTCGTTGTGAACACCCCCC              3334

Qy  1874  GGTGGACCGACTACTGGCAGGCCCACTCGGATCCCCGAGTTCGTGAACACCCCCC          1933
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3335  GGTGGACCGACTACTGGCAGGCCCACTCGGATCCCCGAGTTCGTGAACACCCCCC          3394

Qy  1934  CCCTGGTGAAGCTGCGTGGTACCAGCTGGAGAAGGAGGCCCATCATCGGCGCCGAGACCTTCT  1993
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3395  CCCTGGTGAAGCTGCGTGGTACCAGCTGGAGAAGGAGGCCCATCATCGGCGCCGAGACCTTCT  3454

Qy  1994  ACGTGGACGCGGCGCCGCCAACCGCGAGATCGTGAGCGACGACACCAGAGCTGCAGG       2053
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3455  ACGTGGACGCGGCGCCGCCAACCGCGAGATCGTGAGCGACGACACCAGAGCTGCAGG       3514

Qy  2054  GGGGCCGGCCGAGAGATCGTGAGCGACACCAGGAGACCAGAGCTGCAGG               2113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3515  GGGGCCGGCCGAGAGATCGTGAGCGACACCAGGAGACCAGAGCTGCAGG               3574

Qy  2114  CCATCCAGCTGGCCCTGCGAGGACAGCAGCGGCGGCGAGGTGAACATCGTGACCGACAGCAGT  2173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3575  CCATCCAGCTGGCCCTGCGAGGACAGCAGCGGCGGCGAGGTGAACATCGTGACCGACAGCAGT  3634

Qy  2174  ACGCCCTGGGCATCATCCAGAGGCTGAAGGCCCAGCATCTTCC                     2233
          |||||||||||||||||||||||||||||||||||||||||||
Db  3635  ACGCCCTGGGCATCATCCAGGGCTGAAGGCCCAGCATCTTCC                      3694

Qy  2234  TCATCGAGCAGCTGATCAAGAAGGAGGAAGGTGTACCTGAGCTGGTGGTCGCCCAAGG      2293
```

Left-hand block (positions 74–1094 / 133–2626):

```
Qy   74  AGGGCCCAAGCGCATCATCAAGTGCTTCAACTGGGCAAGAGGAGGGCCACATGCCCGCA    133
Db 1547  AGGGCCCAAGCGCATCATCAAGTGCTTCAACTGCGCAAGGAGGGCCACATGCCCGCA     1606

Qy  134  ACTCCGCGGCCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA  193
Db 1607  ACTCCGCGGCCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA  1666

Qy  194  AGGACTGCACCGAGGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCGAGGGCA   253
Db 1667  AGGACTGCACCGAGGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCGAGGGCA   1726

Qy  254  AGGGCCCGCGAGTTCCCAGCGAGCAGGAACCGCGGCCAACAGCCCGCCACCAGCTGGAGCTGC  313
Db 1727  AGGGCCCGCGAGTTCCCAGCGAGCAGGAACCGCGGCCAACAGCCCGCCACCAGCCGCGGAGCTGC  1786

Qy  314  AGTTGCGGCGGCGACAACCCCGCAGCGAGGCGGCCGCCAGCCGCCAGGCCACCCTGAACT    373
Db 1787  AGTTGCGGCGGCGACAACCCCGCAGCGAGGCGGCCGCCAGCCGCCAGGGCCACCCTGAACT    1846

Qy  374  TCCCCAGATCACCCGTGTGCAGCCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA     433
Db 1847  TCCCCAGATCACCCGTGTGCAGCCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA     1906

Qy  434  AGGAGGCCCCTGCTGACACCGGCCGCGACACCGGGCATGGAGGATGAGCCTGCCCG     493
Db 1907  AGGAGGCCCCTGCTGACACCGGCCGCGACACCGGGCATGGAGGATGAGCCTGCCCG     1966

Qy  494  GCAAGTGGAAGCCCAAGATCGGCGGCGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG   553
Db 1967  GCAAGTGGAAGCCCAAGATCGGCGGCGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG   2026

Qy  554  ACCAGATCCTGATCGAGATCTGCGGCAAGAAGCCATCGGCACCGTGCTCGATCGGCCCCA   613
Db 2027  ACCAGATCCTGATCGAGATCTGCGGCAAGAAGCCATCGGCACCGTGCTCGATCGGCCCCA   2086

Qy  614  CCCCCGTGAACATCATCGGCCGCCAACATGCTGACCTGGCTGGCGCTGCGCACCCTGAACTTCC   673
Db 2087  CCCCCGTGAACATCATCGGCCGCCAACATGCTGACCTGGCTGGCGCTGCGCACCCTGAACTTCC   2146

Qy  674  CCATCAGCCCGCTGACCTGCCGTGAAGCTGAAGCCCGGATGACGGCCCCAAGG          733
Db 2147  CCATCAGCCCGCTGACCTGCCGTGAAGCTGAAGCCCGGATGACGGCCCCAAGG          2206

Qy  734  TGAAGCAGTGGCCCTGACCCTGACCCTGACCAAGGCCCCTGACCGCATCTGCGAGAGA     793
Db 2207  TGAAGCAGTGGCCCTGACCCTGACCCTGACCAAGGCCCCTGACCGCATCTGCGAGAGGA     2266

Qy  794  TGGAGAAGGAGGGCCAAGATCACCAAGGGCCCCGAGAACCCCTACAACACCCCCG        853
Db 2267  TGGAGAAGGAGGGCCAAGATCACCAAGGGCCCCGAGAACCCCTACAACACCCCCGT       2326

Qy  854  TCGGCCATCAAGAAGGACACAGCACCAGTGGGCACCGTGGTGACTTCCGCCAGCTGA      913
Db 2327  TCGGCCATCAAGAAGGACACAGCACCAGTGGGCACCGTGGTGACTTCCGCCAGCTGA      2386

Qy  914  ACAAGGCGCACCCAGGAGCGTGACCCCAGGAGGTGCAGTGGGACGTGGGCATCCCCCTGA   973
Db 2387  ACAAGGCGCACCCAGGAGCGTGACCCCAGGAGGTGCAGTGGGACGTGGGCATCCCCCTGA   2446

Qy  974  AGAAGAAGAAGAGAGTTCCGCGAAGATGGCTGACACGGCTGCGACTTCCGCAAGTGACTTCCG  1033
Db 2447  AGAAGAAGAAGAGAGTTCCGCGAAGATGGCTGACACGGCTGCGACTTCCGCAAGTGACTTCCG  2506

Qy 1034  ACGAGGCGCTTCCGCGAACTGCCGCAGCATCAACAACGAGACCCCCG                1093
Db 2507  ACGAGGCGCTTCCGCGAACTGCCGCAGCATCAACAACGAGACCCCCG                2566

Qy 1094  GCATCCGGTACCAGTACAGCTGCTGCCCCAGGGCTGGAAGGGCCCAGCATCTTCC        1153
Db 2567  GCATCCGGTACCAGTACAGCTGCTGCCCCAGGGCTGGAAGGGCCCAGCATCTTCC        2626
```

```
Best Local Similarity   99.2%;   Pred. No. 1.2e-288;
Matches 2430;  Conservative   0;  Mismatches   8;  Indels  12;  Gaps   2;

Qy      14 TGGCCGAGCCATGAGCCAGGCCAGCCGCCAACATCCTGATGCAGCAACTTCA    73
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    2741 TCGCCGAGCCATGAGCCAGGCCAGCCGCCAACATCCTGATGCAGCGCAACTTCA 2800

Qy      74 AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGCAAGGAGGGGCCACATCGCCCCGCA   133
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2801 AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCCGCA 2860

Qy     134 ACTGCGCGGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGAGTGGGCCACCAGATGA   193
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2861 ACTGCGCGGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGAGTGGGCCACCAGATGA 2920

Qy     194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA   253
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2921 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGGCA 2980

Qy     254 AGGCCCGCGGAGTTCCCCAGCGCGAGCAGAACCGCGCCACAGCCCCAGCCCCGCGAGCTGC   313
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2981 AGGCCCGCGGAGTTCCCCAGCGCGAGCAGAACCGCGCCACAGCCCCAGCCCCGCGAGCTGC 3040

Qy     314 AGGTGCGCGGCGACAACCCCGAGCGCGAGGCCGGCCAGGGCACCCTGAACT   373
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3041 AGGTGCGCGGCGACAACCCCGAGCGCGAGGCCGGCCAGGGCACCCTGAACT 3100

Qy     374 TCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCATCAAGGTGGGCAGGATCA   433
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3101 TCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGCATCAGGATCA 3160

Qy     434 AGGAGGCCCTGCTGGACACCGGGCGCGACGACACCGTGCTGGAGGATGAGCCTGCGCCG   493
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3161 AGGAGGCCCTGCTGGACCACCGGGCGCGACGACACCGTGCTGGAGGATGAGCCTGCGCCG 3220

Qy     494 GCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGGTGCGGCCAGTACG   553
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3221 GCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGGTGCGGCCAGTACG 3280

Qy     554 ACCAGATCCTGATCGGATCTGCGGCCAAGAAGGCCATCGGCCGTGATCGGCCCCA   613
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3281 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGGCCATCGGCCGTGATCGGCCCCCA 3340

Qy     614 CCCCCGTGAACATCATCGGCGCCGCCAGACGTGACCCCGAGACTTCC   673
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3341 CCCCCGTGAACATCATCGGCGCCGCCAGCAGTGACCCCGAGACTTCC 3400

Qy     674 CCATCAGCCCCATCGAGACCCTGCCCGTGAAGTGAAGCCCGGCGGCCCCAAGG   733
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3401 CCATCAGCCCCATCGAGACCCTGCCCGTGAAGTGAAGCCCGGCGGCCCCAAGG 3460

Qy     734 TGAAGCCAGTGGCCCTGACCCTGACGGAGAAGATCAAGGCCCATCGCGAGGAGA   793
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3461 TGAAGCCAGTGGCCCTGACCCTGACGGAGAAGATCAAGGCCCATCGCGAGGAGA 3520

Qy     794 TGGGAAGAGGGGCCAAGATCACCAAGATCGGCGCCCCGAGAACCCCTACAACACCCCGTGT   853
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3521 TGGGAAGAGGGGCCAAGATCACCAAGATCGGCGCCCCGAGAACCCCTACACACACCCGTGT 3580

Qy     854 TCGCCATCAAGAAGGACAGCACCAAGTGGGCGAAGCTGGTGGGACTTGGACTGA   913
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3581 TCGCCATCAAGAAGGACAGCACCAAGTGGGCGAAGCTGGTGGGACTTGGGACTGA 3640

Qy     914 ACAAGGCGCACCCAGGACTTCGGAGGTGCAGTGCGCCCCCACCCCGCGGCCTGA   973
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3641 ACAAGGCGCACCCAGGACTTCGGGGTGCAGTGCGCCCCACCCCGGCCGCCTGA 3700

Qy     974 AGAAGAAGAAGAGCGTGACCGTGCTGACGGTGGACTGGACTGA   1033
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3701 AGAAGAAGAAGAGCGTGACCGTGCTGACGGTGGGACGTGGACTGA 3760

Qy    1034 ACGAGGGACTTCCGGAACGTACACCGCCTTCACCATTCCCCAGCATCAACACGAGACCCCTGG   1093
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3695 TCATCGGACCAGCTGATCAAGAAGCAGAGCAGTGATCGACCAGCAGGGCATCAGGGGCAGCAGGATCGACAAGGAGGAGCAGAGGTGTACCTGAGCTGAGCAGGCAGGCAGCAGGGGCATCCGATGCAGCAAGGTGCTGT 3754
```

```
Db    2294 GCATCGCGGCGACAGCAGAGCAGATCGACAAGCAAGCAAGGTGGTGACGAGCTGGTGAGCAAGGGGCATCCGCAAGGTGCTGT 2353
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3755 GCATCGCGGCGACAGCAGAGCAGATCGACAAGCAAGCAAGGTGGTGACGAGCTGGTGAGCAAGGGGCATCCGCAAGGTGTGT 3814

Qy    2354 TCCTGAGCGACCATGGCGACGATGGGCGCGCATCGTGATCTACCAGTACATGACGACCTGTACGTGG 2413
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3815 TCCTGAGCGACCATGGCGACGATGGGCGCGCATCGTGATCTACCAGTACATGACGACCTGTACGTGG 3874

Qy    2414 GCAGCGCGCGGCCCTAGGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT 2463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3875 GCAGCGCGCGGCCCTAGGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT 3924
```

RESULT 13
ACA03591
ID   ACA03591 standard; DNA; 5184 BP.
XX
AC   ACA03591;
XX
DT   22-MAY-2003 (first entry)
XX
DE   Synthetic DNA ecoding immunogenic HIV peptide #74.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
PT   derived from different HIV subtypes. The polynucleotide is useful for
PT   immunisation, generation of packaging cell lines, or production of HIV
PT   polypeptides. The polynucleotide and its encoded proteins are useful as
PT   immunogenic compositions or vaccines for generating humoral or cellular
PT   immune responses against HIV in a subject, or for inducing neutralising
PT   antibodies against HIV. The gene delivery vector comprising the
PT   polynucleotide is also useful for DNA immunisation of, or for generating
PT   an immune response (e.g. a humoral or cellular immune response) in, a
PT   subject such as a mammal, particularly a human. This sequence encodes a
PT   human immunodeficiency virus immunogenic peptide
XX
PS   Example 1; Fig 79; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, useful as immunogenic compositions or
CC   vaccines for generating humoral or cellular immune responses against HIV
CC   in a subject, particularly a human, in a gene delivery vector for
CC   generating humoral or cellular immune responses against HIV in a subject,
CC   especially humans.
XX
SQ   Sequence 5184 BP; 1139 A; 1852 C; 1610 G; 583 T; 0 U; 0 Other;

Query Match   96.9%; Score 2393.2; DB 7; Length 5184;

Db　　3761 AGAGGACTTCCGCAAGTACACCGCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 3820

Qy　　1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 1153
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　3821 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 3880

Qy　　1154 AGAGCAGCAGCCATGACCAAGATCGTCGGAGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCT 1213
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　3881 AGAGCAGCAGCCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCT 3940

Qy　　1214 ACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGGCAGCACCGCGCCA 1273
　　　　　　 ||| ||　 |||||||||||||||||||||||||||||||||||||||||||||
Db　　3941 ACCA------GGCCCCCTGTACGTGGGGCAGCGACCTGGAGATCGGGCAGCACCGCGCCA 3994

Qy　　1274 AGATCGAGGAGCTGCGCAAGCACCTGCTGCGGCTTCACCACCCCGACAAGAAGC 1333
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　3995 AGATCGAGGAGCTGCGCAAGCACCTGCTGCGGCTTCACCACCCCGACAAGAAGC 4054

Qy　　1334 ACCAGAAGGAGGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCG 1393
　　　　　　 |||||||||||||||||||||||　　　　　||||||||||||||||||||||||||
Db　　4055 ACCAGAAGGAGCCCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACACAAGTGGACCG 4108

Qy　　1394 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGG 1453
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4109 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGG 4168

Qy　　1454 TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA 1513
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4169 TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA 4228

Qy　　1514 AGTTGCTGCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGC 1573
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4229 AGTTGCTGCGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGC 4288

Qy　　1574 TGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACC 1633
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4289 TGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACC 4348

Qy　　1634 CCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGCAGGCCACCGTGG 1693
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4349 CCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGCAGGCCACCGTGG 4408

Qy　　1694 TCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGGCCAAGATGGCACCGCCC 1753
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4409 TCTACCAGGAGACCCTTCAAGAACCTGAAGACCGGCAAGTACGGCCAAGATGGCACCGCCC 4468

Qy　　1754 ACACCAAGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCAGGAGAGCATCG 1813
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4469 ACACCAAGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCAGGAGAGCATCG 4528

Qy　　1814 TGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCT 1873
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4529 TGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCT 4588

Qy　　1874 GGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCCCC 1933
　　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4589 GGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCCCC 4648

Qy　　1934 CCCTGGTGAAGCTGTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 1993
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4649 CCCTGGTGAAGCTGTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 4708

Qy　　1994 ACGTGGACGGCGCCGCCAACCGCGAGATCGGCAAGATCGGCAAGGCCGGCTACGTGACCGACC 2053
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4709 ACGTGGACGGCGCCGCCAACCGCGAGACCGGCAAGATCGGCAAGGCCGGCTACGTGACCGACC 4768

Qy　　2054 GGGGCCGCGCAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGG 2113
　　　　　　 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4769 GGGGCCGCGCAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGG 4828

Qy　　2114 CCATTCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGT 2173
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4829 CCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGT 4888

Qy　　2174 ACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGA 2233
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4889 ACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGA 4948

Qy　　2234 TCATGGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 2293
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　4949 TCATGGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 5008

Qy　　2294 GCATGGCGGCGCAAGGAGCAGATCGACAAGCTGGTGAGCAAGGCATCGCCAAGTGCTGT 2353
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　5009 GCATCGGCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGCATCGCCAAGTGCTGT 5068

Qy　　2354 TCCTGGACGGCATCGATGGCGGGCATCGTGATCTGCACCAGTACATGGACGACCTGTACGTGG 2413
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　5069 TCCTGGACGGCATCGATGGCGGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG 5128

Qy　　2414 GCAGGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT 2463
　　　　　　 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db　　5129 GCAGCCGCGGCCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT 5178

```
RESULT 14
ADC13279
ID   ADC13279 standard; DNA; 5184 BP.
XX
AC   ADC13279;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct TatRevNefgagCpolina_C SEQ ID NO 58.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 55; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
```

```
SQ   Sequence 5184 BP; 1139 A; 1852 C; 1610 G; 583 T; 0 U; 0 Other;

     Query Match       96.9%;    Score 2393.2;  DB 9;  Length 5184;
     Best Local Similarity 99.2%;    Pred. No. 1.2e-288;
     Matches 2430;  Conservative   0;  Mismatches   0;  Indels  12;  Gaps   2;
```

```
Qy     14 TGGCCGAGGCCATGAGCGCCAGGCCACCAGCCCACATCCTGATGCAGCCAGCAACTTCA  73
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 TGGCCGAGGCCATGAGCGCCAGGCCACCAGCCCACATCCTGATGCAGCCAGCAACTTCC 2800

Qy     74 AGGGCCCAGGCCATCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA 133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AGGGCCCCAGGCCATCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA 2860

Qy    134 ACTGCGGCGCCCCCCGCCAAGAGAGGGCTGCTGGAAGTGCGGCAAGGAGTGCGGCAAGGAGTGA 193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 ACTGCGGCGCCCCCCGCCAAGAGAGGGCTGCTGGAAGTGCGGCAAGGAGTGCGGCAAGGAGTGA 2920

Qy    194 AGGATGCACCGAGCGCCAGCCAACTTCTTCCGCGAGAACCGGCCCTTCCCCCAGGGCA 253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGATGCACCGAGCGCCAGCCAACTTCTTCCGCGAGAACCGGCCCTTCCCCCAGGGCA 2980

Qy    254 AGGCCCGGAGTTCCCCAGGCAGCGAGAACCGGCGCGCCACAGCCCCACCAGCCGCGAGCTGC 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 AGGCCCCGGAGTTCCCCAGGCAGCGAGAACCGGCGCGCCACAGCCCCACCAGCCGCGAGCTGC 3040

Qy    314 AGGTGCGGCGGACAACCCCTGTGGCAGCGGCGCCCCCCTGGTGAGCATCAAGTGGGGCGG 373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGGCGGGACAACCCCTGTGGCAGCGGCGCCCCCCTGGTGAGCATCAAGTGGGGCGG 3100

Qy    374 TCCCCCAGATCACCCTGTGGCACCGGCGCCGACGCACACCCCTGTGGAGGAGATGAGCTGCCCG 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCCAGATCACCCTGTGGCACCGGCGCCGACGCACACCCCTGTGGAGGAGATGAGCTGCCCG 3160

Qy    434 AGGAGGCCCTGCTGCAGCACCGGCGCCGACGCACCGGCCCTGCTGCGACACGCTGCCCG 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGCCCTGCTGCAGCACCGGCGCCGACGCACCGGCCCTGCTGCGACACGCTGCCCG 3220

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGGCGGCATCGGCGCCTTCATCAAGGTGCGCCAGTACG 553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATCGGGCGGCATCGGCGCCTTCATCAAGGTGCGCCAGTACG 3280

Qy    554 ACCAGATCCTGATGAGATCTGCGGCAAGAGGCCATGCGCAAGAGGGCTGCTGATCGGCACCCG 613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATGAGATCTGCGGCAAGAGGGCCATGCGCAAGAGGGCTGCTGATCGGCACCCA 3340

Qy    614 CCCCGTGAACATCATCGGCCGCAACATGTCGTGGGCTGCACCCTGAACTTCC 673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCGTGAACATCATCGGCCGCAACATGTCGTGGGCTGCACCCTGAACTTCC 3400

Qy    674 CCATCAGCCCCATCGAGACGTGCCCGTGCCCGTGAAGCTGAAGCTGAAGAGGCCCAAGG 733
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAGACGTGCCCGTGCCCGTGAAGCTGAAGCTGAAGAGGCCCCAAGG 3460

Qy    734 TGAAGCAGTGGCCGCGCCGCCGACTTCTTGGGAGTGCGCTGCAGCTGGCACTCGCGAGGAGA 793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGGCCGCGCCGCCGACTTCTTGGGAGTGCGCTGCAGCTGGCACTCGCGAGGAGA 3520

Qy    794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGT 853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCCGTGT 3580

Qy    854 TCGCCATCAAGAAGAGGAAGGACACAGCACCAAGTGCCACCCAAGTGCTGGTGTGGACTGA 913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCAAGAAGAGGAAGGACACAGCACCAAGTGCCACCCAAGTGCTGGTGTGGACTGA 3640

Qy    914 ACAAGGGCACCCGGACTTCTGGGAGTGCAGCTGCAGTCGCATCCCCACCGGCGCGCCTGA 973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGGGCACCCGGACTTCTGGGAGTGCAGCTGCAGTCGCATCCCCACCGGCGCGCCTGA 3700

Qy    974 AGAAGAAGAAGAGCGTGACCGTGCTGGACGCGTCAGCGTGCTGACGTCAGGTTCAGCGTGCCCTGG 1033
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAAGAAGAAGAGCGTGACCGTGCTGGACGCGTCAGCGTGCTGACGTCAGGTTCAGCGTGCCCTGG 3760
```

```
Qy   1034 ACGAGGACTTCCGCAAGTACACACCGGCTTCACCATCCCCAGCATCAACAACGAGACCCCG 1093
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACGAGGACTTCCGCAAGTACACACCGGCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 3820

Qy   1094 GCATCCGTACCAGTACAACGTGCTGCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC 1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGTACCAGTACAACGTGCTGCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC 3880

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCCTTCCGCGCCCGCAACCCCGAGATCGTGATCT 1213
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3881 AGAGCAGCATGACCAAGATCCTGGAGCCTTCCGCGCCCGCAACCCCGAGATCGTGATCT 3940

Qy   1214 ACCAGTACATGGACGACCTGTACGTGGGCAGCGCAAGGAGGGGCACCAGATGA 1273
          |||||||||||||   ||--------||||||||||||||||||||||||||||
Db   3941 ACCA------GGCCCCTGTACGTGGGCAGCGCAAGGAGGGGCAGCACCAGCCGCCA 3994

Qy   1274 AGATCGAGGAGTGCCAAGCACCTGCTGCCCGGGCTTCACCACCCCGACAAGAAGC 1333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3995 AGATCGAGGAGTGCGCGCAAGCACCTGCTGCCCGGGCTTCACCACCCCGACAAGAAGC 4054

Qy   1334 ACCAGAAGGAGCGCCCCTTCGTGTGGATGGGCTACGAGCTGGTACGAGTGGACCG 1393
          |||||||||||||||||||||||||||||||||      ||||||||||||||||||
Db   4055 ACCAGAAGGAGCGCCCCTTCCTGTGGAACT------CGAGCTGACCCGGACAGTGGACCG 4108

Qy   1394 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGTGGACCGTGAACGACATCCAGAAGCTGG 1453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4109 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGTGGACCGTGAACGACATCCAGAAGCTGG 4168

Qy   1454 TGGGCAAGCTGACTGGGCGCCAAGGGCCTGACCCGGCATCTACCCCGGCCAGCTGTGCA 1513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4169 TGGGCAAGCTGACTGGGCGCCAAGGGCCTGACCCGGCATCTACCCCGGCCAGCTGTGCA 4228

Qy   1514 AGCTGCTGCGCGGCGCCGACACCGGCGCCGACCCTGCCCCTGACCCGGAGGAGGCCGAGC 1573
          |||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db   4229 AGCTGCTGCGCGGCGCCGACACCATCGTGCCCCTGACCCGGAGGAGGCCGAGC 4288

Qy   1574 TGGAGCTGGCCGAGAACCCCGAGATCTTCGCCGCACGGGCCGTGCACCGCGGCGTGCTACTACACC 1633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4289 TGGAGCTGGCCGAGAACCCCGAGATCTTCGCCGCACGGGCCGTGCACCGCGGCGTGCTACTACGACC 4348

Qy   1634 CCAGCAAGGACCTGGTGGCCAAGATCCAGAATCCAGAAGAAGCAGAGGACCGGCACCGCGCC 1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4349 CCAGCAAGGACCTGGTGGCCAAGATCCAGAATCCAGAAGAAGCAGAGGCCACCAGTGGACCTACCACAGA 4408

Qy   1694 TCTACCAGGAGCCCTTCAAGAACTGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4409 TCTACCAGGAGCCCTTCAAGAACTGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 4468

Qy   1754 ACACCAACGACGTGAAGCAGCTGACCGAGAAGCTCGACAACGTGAAGATCGCCATGAGAGCATCG 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4469 ACACCAACGACGTGAAGCAGCTGACCGAGAAGCTCGACAACGTGAAGATCGCCATGAGAGCATCG 4528

Qy   1814 TGATCTGGGGCAAGACCACCCCAAGTTCCGCCTGACCCATCCTGCCCATCGAGGAGACCT 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4529 TGATCTGGGGCAAGACCACCCCAAGTTCCGCCTGACCCATCCTGCCCATCGAGGAGACCT 4588

Qy   1874 GGTGACCGACTACTGGCAGGCCACCTGGACTGGGAGTTCGTGACAACACCCCCC 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4589 GGTGACCGACTACTGGCAGGCCACCTGGACTGGGAGTTCGTGACAACACCCCCCC 4648

Qy   1934 CCCTGGTGAAGCTGTGGTACCGACTGGAGAAGAGCCCATCATCGGCGCCGAGACTTCT 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4649 CCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGAGCCCATCATCGGCGCCGAGACCTTCT 4708

Qy   1994 ACGTGACGCGCGCCCACCCGCGAGAACCGGCCAAGATCGGCAAGGCCGGCTACGTGACCACC 2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4709 ACGTGACGCGCGCCCACCCGCGAGAACCGGCCAAGATCGGCAAGGCCGGCTACGTGACCACC 4768

Qy   2054 GGGCCCGGCAGAAGATCGTGACCGTGAGCCTGGACGCCTGCAGGG 2113
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   4769 GGGCCCGGCAGAAGATCGTGACCGTGAGCCTGGACGCCTGCAGG 4828
```

```
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3531 BP; 802 A; 1210 C; 1096 G; 423 T; 0 U; 0 Other;

Query Match             95.7%;  Score 2362.8;  DB 9;  Length 3531;
Best Local Similarity   98.4%;  Pred. No. 7.3e-285;
Matches 2411;  Conservative    0;  Mismatches   27;  Indels   12;  Gaps    2;

Qy         14 TGGCCCAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACTTCA 73
              |||||||||| ||||||||||||||| |||||||| |||| |||||||||||||||||
Db       1088 TGGCCCAGGCCGATGAGCCAGGCCAACCACCAGCGTGATGATGCAGCGCCAGCAACTTAAAA 1147

Qy         74 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCCA 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1148 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCCA 1207

Qy        134 ACTGCGCGCCCCCGCCAAGAGGGCTGCTGGAAGTGCGGCAAGAGTGCGGCACCAGATGA 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1208 ACTGCGCGCCCCCGCCAAGAGGGCTGCTGGAAGTGCGGCAAGAGTGCGGCACCAGATGA 1267

Qy        194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGAACCGGCCAACAGCCCCCAGGGGCA 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1268 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGAACCGGCCAACAGCCCCCAGGGGCA 1327

Qy        254 AGGCCCGGAGTTCCCAGGCAGCAGAGAACCGGCAACAGCCGCCGAGCGCCGCGAGCTGC 313
              |||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db       1328 AGGCCCGGAGTTCCCAGGCAGCAGAGAACCGGCAACAGCCCCAGCGGCCGCGAGCTGC 1387

Qy        314 AGGTGCGGCGGGACAACCCCCGCGAGGCCGGCCAGGCCGCAGGCCACCCTGAACT 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1388 AGGTGCGGCGGGACAACCCCCGCGAGGCCGGCCAGGCCGCAGGCCACCCTGAACT 1447

Qy        374 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGGGCCAGATCA 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1448 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGGGCCAGATCA 1507

Qy        434 AGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG 493
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1508 AGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG 1567

Qy        494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCGCCAGTACG 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1568 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCGCCAGTACG 1627

Qy        554 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGAGGCCATCGGCGTGCTGGTGATCGGCCCCA 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db       1628 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGAGGCCATCGGCGTGCTGGTGATCGGCCCCA 1687

Qy        614 CCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCGTGCACCCTGAACTTCC 673
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1688 CCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCGTGCACCCTGAACTTCC 1747

Qy        674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 733
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1748 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 1807

Qy        734 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 793
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1808 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 1867

Qy        794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 853
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1868 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 1927

Qy        854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1928 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 1987

Qy        914 ACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 973
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1988 ACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 2047

Qy       2114 CCATCCAGTGGCCCTGCAGGACGAGCGGCGAGCCGAGGTGAACATCGTGACCGACAGCCAGT 2173
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4829 CCATCCAGCTGGCCCTGCAGGACGAGCGGCGAGCCGAGGTGAACATCGTGACCGACAGCCAGT 4888

Qy       2174 ACGCCCTGGGCATCATCCAGGCCCGACAAGCCGAGGAGCGAGCTGGTGAACCAGA 2233
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4889 ACGCCCTGGGCATCATCCAGGCCCGACAAGCCGAGGAGCGAGCTGGTGAACCAGA 4948

Qy       2234 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 2293
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4949 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 5008

Qy       2294 GCATCGGCGGCCAACGAGCCAGATCGACAAGCTGGTGAGCAAGGGCCATCCGCAAGGTGCTGT 2353
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       5009 GCATCGGCGGCCAACGAGCGAGCAGATCGACAAGCTGGTGAGCAAGGGCCATCCGCAAGGTGCTGT 5068

Qy       2354 TCCTGGACGGCATCGATGGCGGCATCGTGATCCACCAGTACACATGGACGACCTGTACGTGG 2413
              ||||||||||||||||||||||||||||| |||||||||| |||||||||||||||||||||
Db       5069 TCCTGGACGGCATCGATGGCGGCATCGTGATCGTGATCCACCAGTACACCAGTACATGGACGACCTGTACGTGG 5128

Qy       2414 GCAGCGGCGGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2463
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5129 GCAGCGGCGGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 5178
```

RESULT 15
ADC13234
ID   ADC13234 standard; DNA; 3531 BP.
XX
AC   ADC13234;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct GagPolmut_C SEQ ID NO 13.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 10; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV

```
Qy   974 AGAAGAAGAAGAGAGCGTGACCGTGCTGGGCGACGCCTACTTCAGCGTGCCCCTGG 1033
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2048 AGAAGAAGAAGAGAGCGTGACCGTGCTGGGCGACGCCTACTTCAGCGTGCCCCTGG 2107

Qy  1034 ACGAGGACTTCCGCAAGTACACCGGCCTTCACCATCCCAGCATCAACAACGAGACCCCCG 1093
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2108 ACGAGGACTTCCGCAAGTACACCGGCCTTCACCATCCCAGCATCAACAACGAGACCCCCG 2167

Qy  1094 GCATCCGCTACCAGTACAAACGTGCTGCCCCAGGGGTGGAAGGGCAGCCCCCAGCATCTTCC 1153
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2168 GCATCCCGCTACCAGTACAAACGTGCTGCGTCCCCAGGGCGTGGAAGGGCAGCCCCCAGCATCTTCC 2227

Qy  1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCCCCGCCAACCCCGAGATCGTGATCT 1213
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2228 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCCAACCCCGAGATCGTGATCT 2287

Qy  1214 ACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCA 1273
         |||||||||||||                       |||||||||||||||||||||||
Db  2288 ACCA------GGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCA 2341

Qy  1274 AGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCACAAGAAGC 1333
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2342 AGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCACAAGAAGC 2401

Qy  1334 ACCAGAAGGAGCGCCCCCTTCCTGTGGATGGGCTACGAGCTGACACCCGACACAAGTGACCG 1393
         ||||||||||||||||||||              ||
Db  2402 ACCAGAAGGAGCGCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACACAAGTGACCG 2455

Qy  1394 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGACCTGAACGGTGAACGACATCCAGAAGCTGG 1453
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2456 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGAGCTGGACCGTGAACGACATCCAGAAGCTGG 2515

Qy  1454 TGGGCAAGTGAACTGGGCCGCGGATCTACCCGGCGCATCAAGGTGCGCCCAGCTGTGCA 1513
         |||||||||||||||||||      ||||||||||||||||||||||||||||||||||
Db  2516 TGGGCAAGTGAACTGGGCGCGGCCAGCCAGATCTACCCCGGCGCATCAAGGTGCGCCCAGCTGTGCA 2575

Qy  1514 AGCTGCTGCGCGGCGCCAAGGCCCTGACCCCTGACCCGAGGAGGGCGAGC 1573
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2576 AGCTGCTGCGCGGCGCCAAGGCCCTGACCCCTGACCCGAGGAGGGCGAGC 2635

Qy  1574 TGGAGCTGGCCGGAGAACCGGCGAGATCCTTGCGGCGCCCGTGCACGGCGTGTACTACGACC 1633
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2636 TGGAGCTGGCCGGAGAACCGGCGAGATCCTTGCGGCGCCCGTGCACGGCGTGTACTACGACC 2695

Qy  1634 CCAGCAAGGACCTGGTGCGCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGA 1693
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2696 CCAGCAAGGACCTGGTGCGCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGA 2755

Qy  1694 TCTACCAGGAGCCCTTCAAGAACCTCAAGAACCTCAAGACCCGGAAGACCGGCAAGTGGCACCGCGCCC 1753
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2756 TCTACCAGGAGCCCTTCAAGAACCTCAAGACCCGGAAGACCGGCAAGTGGCACCGCGCCC 2815

Qy  1754 ACACCAACGACGTGAAGCAGCTGACCAGCTGAGAAGAGACCCGTGCGATCGGCCGATCG 1813
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2816 ACACCAACGACGTGAAGCAGCTGACCAGCTGAGAAGAGACCCCATCATCGGCGCCGATCG 2875

Qy  1814 TGATCTGGGGCAAGACCCCCAGTTCGGCCTGCCGCCCATCCAGAAGGAGACCTGGGAGACCT 1873
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2876 TGATCTGGGGCAAGACCCCCAGTTCGGCCTGCCGCCCATCCAGAAGGAGACCTGGGAGACCT 2935

Qy  1874 GGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTCGTGAACACACCCCC 1933
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2936 GGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTCGTGAACACACCCCC 2995

Qy  1934 CCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 1993
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2996 CCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 3055

Qy  1994 ACGTGGACGGCGCGCCAACCCGCGAGACCCAAGATCGGCCAAGGCGCCGGCTACGTGACCGACC 2053
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3056 ACGTGGACGGCGCGCCAACCGCGAGACCCAAGATCGGCCAAGGCGCCGGCTACGTGACCGACC 3115

Qy  2054 GGGGCCGCGGCAGAAGATCGTGAGCCTGACCGAGACCAGAGACCGAGCTGCAGG 2113
```

```
Db  3116 GGGGCCGGCGGAGAGATCGTGAGCCTGACCGAGACCAGAGACCGAGCTGCAGG 3175
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2114 CCATCCAGCTGGCCCTGCGAGGACGCGCAGCGGCAGGTGAACATCGTGACCGACAGCCAGT 2173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3176 CCATCCAGCTGGCCCTGCGAGGACGCGCAGCGGCAGGTGAACATCGTGACCGACAGCCAGT 3235

Qy  2174 ACGCCCTGGGCATCATCCAGGCCCAAGAGCGAGAGCGAGCTGGTGAACCAGA 2233
         |||||||||||||||||||||||      ||||||||||||||||||||||||
Db  3236 ACGCCCTGGGCATCATCCAGGCCCAGCCCGGAGAGCGAGAGCGAGCTGGTGAACCAGA 3295

Qy  2234 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCACAAGG 2293
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3296 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCACAAGG 3355

Qy  2294 GCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGT 2353
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3356 GCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGT 3415

Qy  2354 TCCTGGACGGCATCGATGGCGGCATCGTGATCTGATCGTTACCAGTACATGGACGACCTGTACGTGG 2413
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3416 TCCTGGACGGCATCGATGGCGGCATCGTGATCGTGATCTACCAGTACATGGACGACCTGTACGTGG 3475

Qy  2414 GCAGCGGCGCCCTAGGATCGGATTAAAAGCTTCCCGGGCGCTAGCACCGGT 2463
         ||||||||||      |||||||||||||||||||||||||||||||
Db  3476 GCAGCGGCCGCCCTAGGATCGATTAAAAGCTTCCCGGGCGCTAGCACCGGT 3525

Search completed: April 10, 2004, 07:32:20
Job time : 632.513 secs
```

```
                 Copyright (c) 1993 - 2004 Compugen Ltd.
                          GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on:        April 10, 2004, 06:48:32 ; Search time 118.622 Seconds
                                         (without alignments)
                                         11550.795 Million cell updates/sec

Title:         US-09-610-313-30
Perfect score: 2469
Sequence:      1 gtcgacgccaccatggccga.........gggctagcaccggtgaattc 2469

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:   1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
        1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
        2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
        3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
        4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
        5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
        6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2052	83.1	2312	4	US-09-475-515-84	Sequence 84, Appl
2	2025.2	82.0	2306	4	US-09-475-515-82	Sequence 82, Appl
3	1998.4	80.9	2300	4	US-09-475-515-83	Sequence 83, Appl
4	1969.4	79.8	4319	4	US-09-475-515-6	Sequence 6, Appli
5	1858	75.3	2305	4	US-09-475-515-80	Sequence 80, Appl
6	1831.2	74.2	2299	4	US-09-475-515-81	Sequence 81, Appl
7	1678.6	68.0	4307	4	US-09-552-950-2	Sequence 2, Appli
8	1651.4	66.9	9772	4	US-09-552-950-5	Sequence 5, Appli
9	1592.8	64.5	8366	4	US-09-872-733A-6	Sequence 6, Appli
10	1557.4	63.1	4338	4	US-09-872-733A-1	Sequence 1, Appli
11	1227.2	49.7	9010	4	US-09-184-418C-8	Sequence 8, Appli
12	1196.2	48.4	8972	4	US-09-184-418C-9	Sequence 9, Appli
13	1190.6	48.2	2467	4	US-09-872-733A-3	Sequence 3, Appli
14	1189	48.2	8959	4	US-09-184-418C-11	Sequence 11, Appl
15	1165.6	47.2	8992	4	US-09-184-418C-4	Sequence 4, Appli
16	1132	45.8	2601	3	US-09-117-217-7	Sequence 7, Appli
17	1132	45.8	2601	3	US-09-117-217-9	Sequence 9, Appli
18	1132	45.8	2601	3	US-09-117-217-11	Sequence 11, Appl
19	1132	45.8	2601	3	US-09-117-217-13	Sequence 13, Appl
20	1132	45.8	2601	4	US-09-735-487-7	Sequence 7, Appli
21	1132	45.8	2601	4	US-09-735-487-9	Sequence 9, Appli
22	1132	45.8	2601	4	US-09-735-487-11	Sequence 11, Appl
23	1132	45.8	2601	4	US-09-735-487-13	Sequence 13, Appl
24	1132	45.8	4307	4	US-09-552-950-1	Sequence 1, Appli
25	1132	45.8	9719	4	US-09-700-304-1	Sequence 1, Appli
26	1128.8	45.7	9050	4	US-09-184-418C-7	Sequence 7, Appli
27	1125.6	45.6	7399	2	US-08-418-848A-9	Sequence 9, Appli
28	1125.6	45.6	9709	2	US-08-188-583-5	Sequence 5, Appli
29	1125.6	45.6	9709	3	US-08-388-353-1	Sequence 1, Appli
30	1125.6	45.6	9709	3	US-08-488-555B-1	Sequence 15, Appl
31	1125.6	45.6	9709	4	US-09-309-572-15	Sequence 15, Appl
32	1125.6	45.6	9709	4	US-09-718-096-15	Sequence 13, Appl
33	1125.6	45.6	12479	4	US-09-318-138-13	Sequence 13, Appl
34	1125.6	45.6	12494	3	US-08-935-312-13	Sequence 33, Appl
35	1125.6	45.6	12494	3	US-08-848-760B-33	Sequence 35, Appl
36	1125.6	45.6	15581	3	US-08-646-538-35	Sequence 35, Appl
37	1125.6	45.6	15581	3	US-09-503-222-35	Sequence 74, Appl
38	1122.8	45.5	3000	4	US-09-184-418C-74	Sequence 1, Appli
39	1121	45.4	8968	4	US-09-184-418C-1	Sequence 3, Appli
40	1116	45.2	4313	4	US-09-475-515-3	Sequence 7, Appli
41	1116	45.2	9737	4	US-08-944-449-7	Sequence 7, Appli
42	1116	45.2	9737	4	US-09-353-362-7	Sequence 6, Appli
43	1108	44.9	8954	4	US-09-184-418C-6	Sequence 83, Appl
44	1101.2	44.6	3017	4	US-09-184-418C-83	Sequence 101, App
45	1099	44.5	3011	4	US-09-184-418C-101	

```
                          ALIGNMENTS

RESULT 1
US-09-475-515-84
; Sequence 84, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 2312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(-).protmod.RTopt(+)
US-09-475-515-84

Query Match             83.1%;  Score 2052;  DB 4;  Length 2312;
Best Local Similarity   93.6%;  Pred. No. 0;
Matches 2165;  Conservative   0;  Mismatches 135;  Indels 12;  Gaps 2;

Qy  170 GCGCGCAAGGAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCG 229
            | ||||| |||||  ||||||||||||| |||||||| |||| |||  | |||||||||||||
Db    1 GCGCGCCGCAAGGACACCAGACACCAAATGAAGATTGCACTGAGACGAGAGGCTAATTCTTCCGCG 60

Qy  230 AGGACCTGGCCTTCCCCAGGGCAAGGCCCGCGAGTTCCCAGCGGAGCAGCAGACCGCGCCA 289
            ||||||||||||||  | ||||||||||||||||||| || ||||||| ||||||||||
Db   61 AGGACCTGGCCTTCTCGCAGGGCAAGGCCCGCGAGTTCAGCAGCAGACCGAGCAGACCCGCGCCA 120

Qy  290 ACAGCCCCACCAGCGCGCGAGCTGCGAGGTGCGCGGCG------ACAACCCCGCAGCGAGG 343
            ||||||||||| |||||||||||||||||||||||||      |||| | ||| || |||
Db  121 ACAGCCCCACCCGCGCGCGAGCTGCGAGGTGCGCGGCGGAGAACAACAGCCGCAGGGAGG 180

Qy  344 CCGGCGCCGAGCGCGCCGCAGGGCACCCTG------AACTTCCCCCAGATCACCGTGGCAGC 397
            |||||||| |  |||  | |||||||| |||      ||||||||||||||||||||||||||
Db  181 CCGGCGCCACCGGCGGGGCACCGTGAGCTTCAACTTCCCCCAGATCACCGTGGCAGC 240
```

```
Qy   398 GCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGAGGCCCTGCTGCTGACACCGGCG 457
         ||||||||||||||  ||||||  |||||||||||||  |||||||  |||||||  |||||||
Db   241 GCCCCCTGGTGACCATCAGGATGCGCGGCGGCCAGCTTCAAGGAGGCGCTGCTGCTGCGACACCGGCG 300

Qy   458 CCGACGACACCGTGCTGGAGAGATGAGCCTGCCCCGCAAGTGAGCCCAAGATGATCG 517
         |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db   301 CCGACGACACCGTGCTGCTGGAGAGATGAGCCTGCCCCGCAAGTGGAAGCCCAAGATGATCG 360

Qy   518 GCGGGCATCGGCGGCGCTTCATCAAGGTGCGCGCAGTACGACCAGATCCTGATCGAGATCTGCG 577
         ||||||  ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db   361 GCGGGGATCGGCGGGGCTTCATCAAGGTGCGCGCAGTACGACCAGATCCCCGTGCGAGATCTGCG 420

Qy   578 GCAAGAAGGCATCGGCACCGTGCTGATCGGCCGCCCCCCGTGAACATCATCGGCCGCA 637
         ||||||||||||||||||||||  ||||||||||||||  |||||||||||||||||
Db   421 GCCACAAGGCCATCGGCACCGTGCTGGGCCCGTGCGCACCCCGTGAACATCATCGGCCGCA 480

Qy   638 ACATGCTGACCCAGTGGGCTGCACCCTGAACTTCCCATCAGCCCATCGAGACCGTGC 697
         |||||||||||||||||||||||  |||||||||||||||||||||||  ||||||||
Db   481 ACCTGCTGACCCAGATGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACGGTGC 540

Qy   698 CCGTGAAGCTGAAGGCCCTGAGCGGACGGGCCCAAGTGAAGCAGTAGCAGTGGAAGATCTGCG 757
         |||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db   541 CCGTGAAGCTGAAGGCCCTGAGCGGGGGATGGACGGGCCCAAGTCAAGCAGTGGCAAGATCAGCA 600

Qy   758 AGAAGATCAAGGCCCTGACCGGCATCTGCGAGGAGATGAGGAGGGGCAAGATCACCA 817
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGAAGATCAAGGCCCTGACCGGCATCTGCGAGGAGATGGGAGAAGGGGCAAGATCAGCA 660

Qy   818 AGATCGGCCCCGAGAACACCCGTGTTCGCCATCAAGAAGGACAGCA 877
         |||||||||||||||||||||||  |||||||||||||||||||||
Db   661 AGATCGGCCCCGAGAACACCCCGTGTTCGCCATCAAGAAGGACAGCA 720

Qy   878 CCAAGTGGCGCAAGCTGGTGGACTTCCGGGACTGGAACAGGCGCACCCAGGACTTCTGGG 937
         ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||
Db   721 CCAAGTGGCGCAAGCTGGTGGATGGACTTCCGGGACTGGAACAGCGCACCCAGGACTTCTGGG 780

Qy   938 AGGTGCAGCTGGGCATCCCCCACCCCGGCGGCCTGAAGAGAAGAGTGCACCGTGCC 997
         ||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db   781 AGGTGCAGCTGGGCATCCCCCCACCCCGGCGGCCTGAAGAGAAGAGTGCACCGTGCC 840

Qy   998 TGGACGTGGGCGACGCGCTACTTCAGCGGACGGGCACCGCTACGAAGTTCCGGGACGTACACCG 1057
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGGACGTGGGCGACGCGCTACTTCAGCGGGCACGCGCTACGAAGTTCCGGGACGTACACCG 900

Qy  1058 CCTTCACCATCCCCAGCATCAACAAGAGACCCCGGCATCCCCGGCATCTACCAGTACAACGTGC 1117
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 CCTTCACCATCCCCCAGCATCAACAAGAGACCCCGGCATCCCCGGCATCTACCAGTACAACGTGC 960

Qy  1118 TGCCCCAGGGTGGAAGGGCAGCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TGCCCCAGGGCTGGAAGGGCAGCGCAGCCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020

Qy  1178 AGCCCTTCCGCGGCAACCCCGAGATCGTGATCTACCAGTACCACCAGGACGACCTGTACG 1237
         |||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db  1021 AGCCCTTCCGCAAGCAGATCACCCGACATGGCAATGCTACCAGTACAACGACGACCTGTACG 1080

Qy  1238 TGGGCAGCGAGCTGGAGATCGGCGCCAGCAGCGCCGCACCAGCTGCGCCAGCACC 1297
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TGGGCAGCGAGCTGGAGATCGGCGCCAGCAGCGCCGACCAGCTGCGCCAGCACC 1140

Qy  1298 TGCTGCGCGTGGGGCTTCACCACCCCCACCACCCCCGACAAGAGCACCACCCCCCCTTCCTGT 1357
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TGCTGCGCGTGGGGCTTCACCACCCCCCACCACCCCCGACAAGAGCACCACCCCCCCTTCCTGT 1200

Qy  1358 GGATGGGCTACAGCAGCTGCACCCGCAAGTGACCGTGCAGTCGCCCATCGAGCTGCCCCGAGA 1417
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GGATGGGCTACAGCAGCTGCACCCGCAAGTGGGCAAGCTGAACTGGCCCCATCATGCTGCCCCGAGA 1260

Qy  1418 AGGAGAGCTGGACCGTGAACGACTCCAGAAGCTGGTGTGGGCAAGCTGAACTGGCCAGCC 1477
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AGGACAGCTGGACCGTGGAACGACATCCAGAGAGCTGGTGTGGGCAAGCTGAACTGGCCAGCC 1320

Qy  1478 AGATCTACCCCGGCATCAAGGTGCGCGCCAGCTGTGCGCAAGCTGCTGCGCGGCGGCCAAGGCCC 1537
```

```
Db  1321 AGATCTACGCCGGCATCCAAGGTGAAGCAGCAGCTGTCGCAAGCTGCTGCGCGGCCACCAAGGCCC 1380

Qy  1538 TGACCGACATCGTGCCCCTGCCCCTGACCGAGTGGCAGCTGTCGGAGTGGAGCTGGCCCGAGA 1597
         ||||||||||| |||||||||||||||||||  |||||||||||  |||||||||||||||||
Db  1381 TGACCGAGTGATCCCCCTGACCGAGGAGGCCCAGCTGGAGCTGGCCCGAGA 1440

Qy  1598 TCCTGCGGCGCCCGTGCACGGCGTGCTTACGAGTGTACGACCCCCAGCAAGAACCTGTGGCCCGAGA 1657
         ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 TCCTGAAGGCGCCGTGCAGAGGCGTGCTTACGACTGTACGACCCCCAGCAAGAACCTGGTGGCCCGAGA 1500

Qy  1658 TCCAGAAGCAGGGCCAGCCAGTGGACCAGCAGGTGGACCTACCAGGAGATCTACCAAGAACC 1717
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 TCCAGAAGCAGGGCCAGCCAGTGGACCAGCAGGTGGACCTACCAGGAGATCTACCAAGAACC 1560

Qy  1718 TGAAGACCGCAAGTACGCCAAGATGCGCCAAGAGCAGACCACCCAGAACGAGCAGCTGA 1777
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 TGAAGACCGGCAAGTACGCCCGCCCGGCGGGCCCCCACCACCAGGAAGAGTGAGCAGCTGA 1620

Qy  1778 CCGAGGCCGTGCAGAGATCGGCCATGGACATCGGTGATCTGGGCAAGAACACCCCAAGT 1837
         |||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db  1621 CCGAGGCCGTGCGCAGAGGCCCTGGGATGCACCGGCAAGATCGTGATCTGGGCAAGATCCCCAAGT 1680

Qy  1838 TCCCCCTGCCATCCAGAAGGAGGAACCTGGGAGACCTGGAGACGTACTGCGAGGCCA 1897
         |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db  1681 TCAAGCTGCCATCCAGAAGGAGGGAGGACCTGGAGACCTGGAGATGGATGACTGCTGCGAGGCCA 1740

Qy  1898 CCTGGATCCCCAGTGGGAGTTCGTGAACACCCCCCCGTGGAACTTCTACGTGACCGGGGGGCCAGC 1957
         |||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||
Db  1741 CCTGATCCCCAGTGGGAGTTCGTGAACACCCCCCGTGGATGACTTCTACGTGGGACGGGCCAGC 1800

Qy  1958 TGGGAAGGAGCCCATCATCGTGGGCGGCCGCCTACGTGGGGCGGCCGCGAAGATCGTGAGCGCC 2017
         |||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db  1801 TGGGAAGGAGCCCATCATCGTGGGCGGCCGCCTACGTGGGGCGGCCGCGAAGATCGTGAGCGCG 1860

Qy  2018 AGACCAAGATCGGCAAGGCCCTAGTGTGACCGACCGTGAGCC 2077
         ||||||||||||||||||||||||||||||||||||||||||
Db  1861 AGACCAAGGTGGGCAAGGCCCTACGTGACCGACCGTGAGCA 1920

Qy  2078 TGACCGACCACCAACCAGAAGACCAGCTGCAGGCCATCAGCTGCCCTGCAGGGACA 2137
         |||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db  1921 TCGCCGACACCACCAACCAGAAGACCAGCTGCAGGCCATCATCCAGCTGCCCTGCAGGGACA 1980

Qy  2138 GCGGCAGCAGCTGGAACATCGTGACGACGCCAGTAGCGCCCTGGGCATCATCCAGGCCC 2197
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 GCGGGCCTGGGAGGTGAACATCGTGACGACGCCAGTAGCGCCCTGGGCATCATCCAGGCCC 2040

Qy  2198 AGCCCCGACAAGAGAGGCGAGCGAGCGAGCTGGTGAACCAGATCGATCGAGCAAGAAGG 2257
         ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db  2041 AGCCCCGACAAGAGAGGCGAGCGAGCGAGCTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGG 2100

Qy  2258 AGAAGTGTACCTGACGCTGGGTGCCCGGTGCCCACAAGGGCATCGGCGCGCAACGAGCAGATCG 2317
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 AGAGGTGTACCTGCCTGGGTGCCCGGTGCCCCACAAGGGCATCGGCGCGCAACGAGCAGGTCG 2160

Qy  2318 ACAAGCTGGTGAGCAAGGGCATCGCCAGCACCTGCATCGATGCGGCGGCA 2377
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 ACAAGCTGGTGAGCAAGGGCATCGCCAGCACCTGCATCGATGCGGCGGCA 2220

Qy  2378 TCGTGATCTACCAGTACATGACGACCTGTACGGTGGGCAGCGGCCCTAGGATGATT 2437
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 TCGTGATCTACCAGTACATGACGACCTGTACGGTGGGCAGCGGCCCTAGGATGATT 2280

Qy  2438 AAAAGCTTCCCGGGGCTAGCACCCGTGAATTC 2469
         ||||||||||||||||||||||||||||||||
Db  2281 AAAAGCTTCCCGGGGGCTAGCACCCGTGAATTC 2312
```

RESULT 2
US-09-475-515-82
; Sequence 82, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:

```
;  APPLICANT: BARNETT, Susan
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: SRIVASTAVA, Indresh
;  APPLICANT: LIAN, Ying
;  APPLICANT: HARTOG, Karin
;  APPLICANT: LIU, Hong
;  APPLICANT: GREER, Catherine
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
US-09-475-515-82

;  SEQ ID NO 82
;  LENGTH: 2306
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: FS(-).protmod.RTopt.YM
```

```
Query Match            82.0%;  Score 2025.2;  DB 4;  Length 2306;
Best Local Similarity  93.3%;  Pred. No. 0;
Matches 2156;  Conservative 0;  Mismatches 138;  Indels 18;  Gaps 3;
```

```
Qy   170  GCGGCAAGGAGGGCCAGAGTGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCG  229
               |||||||||||||| | | |||   || |||   |||||   ||||||||||||||
Db     1  GCGGCCGCGAAGGACACCAAATGAAAGATTGCACTGAGACAGCCTAATTTCTTCCGCG   60

Qy   230  AGGACCTGCCCTTCCCCCGAGGCAAGGCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCA  289
               ||||||||||||| | |||||||  ||  |   ||||||||||||   |||||
Db    61  AGGACCTGGCCTTCTGCTGCAGGGCAAGGGCCCGCGAGTTCAGCGAGCAGCAGACCGAGG  120

Qy   290  ACAGCCCCACCAGCCGCCGCGAGCTGCAGGTCGCGCGCG------ACAACCCCGCAGGAGG  343
          |||||||||||||||||||||||||||||  |  | ||       |  |   || |||||
Db   121  ACAGCCCCACCCGCCGCGAGCTGCAGGTGTGGGCGCGCCGAGAACAACAGCCTGAGCGAGG  180

Qy   344  CCGGCGCCGAGCGGCCAGGGCACCCTG-----AACTTCCCCAGATCACCCGTGGCAGC  397
          |||||||||||||||||||| |||||      |||||||||||||||  |||||||||
Db   181  CCGGCGCCGACCGGCCAGGGCACCCTGCGAGCTTCAACTTCCCCAGATCACCCGTGGCAGC  240

Qy   398  GCCCCCTGGTGAGATCAAGGTGGGCGCGGCAGATCAAGGAGGCCCTGCTGAGCACCCGCG  457
          ||||||||||||||||||||  ||||||||||  ||||  ||||||  |  |||||||||
Db   241  GCCCCCTGGTGACCATCAGGATCGGCGCGGCAGCTCAAGGAGGCCCTGCTCGACACCCGCG  300

Qy   458  CCGAGCACACCGTGGCTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGAAGCCAAGATGATCG  517
          ||||| ||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db   301  CCGACGAGCACCGTGGCTGCTGGAGGAGATGAACCTGCCCGGCAAGTGAAGCCAAGATGATCG  360

Qy   518  GCGGCATCGGCGGCTTCATCAAGGTGCGCGAGTACGACCCAGATCCTGATCGAGATCTGCG  577
          ||||||||||||||||||||||||||||||||||||||| || ||   ||||||||||||
Db   361  GCGGGATCGGCGGCTTCATCAAGGTGCGCGAGTACGACCCAGATCCTGGAGAATCTGCG  420

Qy   578  GCAAGAAGGCCATGGGCGCACCTGCTGATCGGGCCCCACCCCGTGAACATCATCGGGCCGCA  637
          |||||| ||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db   421  GCCACAAGGCCATCGGCGCACCGCCATCGGCGCCCACCCCGTGAACATCATCGGCCGCA  480

Qy   638  ACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGC  697
          ||||||||||||||||||||| |||||| ||||||||||||||||||||||||||||||
Db   481  ACCTGCTGACCCAGATCGGCTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGGTGC  540

Qy   698  CCGTGAAGCTGAAGCCCGGCATGACGGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG  757
          ||||||||||||||||||||||||||  ||||| ||||||| |||||||||||||||||
Db   541  CCGTGAAGCTGAAGCCCGGGATGGACGGCGGGGGGGATGGCAAGGTCAAGCAGTGGCCCCTGACCGAGG  600

Qy   758  AGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCAGCA  817
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   601  AGAAGATCAAGGCCCTGGTGAGATCTGCACCGAGGAGATGGAGAAGGAGGGCAAGATCAGCA  660
```

```
Qy   818  AGATCGGCCCCGAGAACCCGTACAACACCCCGGTGTTCGCCATCAAGAAGAAGGACAGCA  877
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   661  AGATCGGCCCCGAGAACCCTACAACACCCCGGTGTTCGCCATCAAGAAGAAGGACAGCA  720

Qy   878  CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG  937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG  780

Qy   938  AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC  997
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  AGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGC  840

Qy   998  TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGAGTTCCGCAAGTACACCG  1057
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   841  TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGACAAGTTCCGCAAGTACACCG  900

Qy  1058  CCTTCACCATCCCAGCATCAACAACGACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC  1117
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   901  CCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGC  960

Qy  1118  TGCCCCAGGGGTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG  1177
          |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db   961  TGCCCCAGGGGTGGAAGGGCAGCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCCTGG  1020

Qy  1178  AGCCCTTCCGCGCCCAGAACCCGGACGTGTGATCTCACCAGTACATGGACGACCTGTACG  1237
          ||||||||||||||||||||||||| ||||||||||||||         ||||||||||
Db  1021  AGCCCTTCCGCGCCCAAGCAGAACCCGACATCGTGATCTCACCA------GGCCCCTGTACG  1074

Qy  1238  TGGGCAGCGACCTGGAGATCGGCCAGCACCGCGACAAGATCGAGAAGCTGCGCGCAGCACC  1297
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  1075  TGGGCAGCGACCTGGAGATCGGCCAGCACCGCGACAAGATCGAGAAGCTGCGCCAGCACC  1134

Qy  1298  TGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGT  1357
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1135  TGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGT  1194

Qy  1358  GGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGTGCTGCCCGAGA  1417
          |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  1195  GGATGGGCTACCAGCTGCACCCCGACAAGTGGACCGTGCAGCCGCCATCATGCTGCCCGAGA  1254

Qy  1418  AGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCC  1477
          ||||| ||||| |||||||| |||||||||| |||||||||||||||| |||||||||||
Db  1255  AGGACAGCTGGACCGTGAACGACCATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCC  1314

Qy  1478  AGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCCACCGCCCC  1537
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1315  AGATCTACGCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCACCAAGGCCCC  1374

Qy  1538  TGACCGACATCGTGCCCCTGACCGAGGAAGTGCAAGTGCTGAGCGCCGAGAACCGCGAGA  1597
          ||||||||||| ||||||||||||||| |||||||||||||||||||||||||||||||
Db  1375  TGACCGAGGTGATCCCCCTGACCGAGGAAGTGCAAGTGCTGGAGAGCTGGGCCGAGAACCGCGAGA  1434

Qy  1598  TCCTGCGGCGAGCCCGTGCACGGCGTGTACTACGACCCAGAAGGATCCTGATCGAGATCTGCG  1657
          ||||||||||||||||||||||||| |||||||||||||||||| ||||| |||||||||||
Db  1435  TCCTGAAGGAGCCCGTGCACGGCGTGTACTACGACCCAGCAGGATCTGTGGGCGAGATCTGCG  1494

Qy  1658  TCCAGAAGCAGGGCCACGACCAGCTGACCGAGGCCCTTCACCAGAGCCCTTCAGGAACC  1717
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1495  TCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAGGAACC  1554

Qy  1718  TGAAGACCGGCAAGTACGCCAAGATGCGCACCGCGCCCGCCCCACCAACACGACGCTGA  1777
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  1555  TGAAGACCGGCAAGTACGCCAAGATGCGCACCGCGCCCGCCCCACCAACACGACGCTGA  1614

Qy  1778  CCGGTGAAGCTGAGAGATCGGCATGACCCCATCTGGGCGCAAGACCCCAAGT  1837
          ||||||||||||| || |||||| | ||||||||||||||||| |||||||
Db  1615  CCGAGGAGCCCGTGCGTGCACGAGGTGTACTACGAGCATCCACCGGCAAGATCCCCAAGT  1674

Qy  1838  TCCGCCTGCCCATCGACCGCCATCTGCGGAGGAGACCTGGTGGTGACCGAGGCCA  1897
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  1675  TCAAGCTGCCCATCCAGAAGGAGACCTGGGAGGAGGTGGGCAGGCCGCAGCGGCCA  1734

Qy  1898  CCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCTTGTGAAGCTGTGGTGGTACCAGC  1957
```

```
Db   170 GCGGCAAGAGAGGGCACCAGATGAAGGACTGCACCGAGCGCCAGGCAGCCAACTTCTTCCGCG  229
             ||||| ||  ||  |||| |||||| ||||||||||||||| | ||||||||||||||||
Qy     1 GCGGCCGCGGAAGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTCTTCCGCG     60

Qy   230 AGGACCTGGCCTTCCCCAGGGCCAAGCCCGGAGTTCCCAGCGAGACGAGACCGCGCCA    289
             |||||||||| ||| | |||| || ||||||||||| ||||| ||| ||||| | 
Db    61 AGGACCTGGCCTTCCTGCAGGGCGGCCTAGTGACGACCGGGCCGGAGAATCGTGAGCC     120

Db   290 ACAGCCCCACCAGCGCGAGCTGCAGGTGCGCGCG-----ACAACCCCAGCGAGG     343
             ||||||| || |||| ||||||| ||||| ||| 
Qy   121 ACAGCCCCACCCGCGCGAGCTGCAGGTGTGGGGCGGCGAGAACAACAGCTGAGCGAGG    180

Qy   344 CCGGCGCCAGCGGCCAGGCACCCTG-----AACTTCCCCAGATCACCCGTGGCAGC    397
             |||||||||| || |||||| |||    |||||||||||| | |||||||||||| 
Db   181 CCGGCGCCCGACCCGCACCGGGCCGCTTCAACTTCCCCAGATCACCCGTGGCAGC     240

Db   398 GCCCCCTGGTGACGCAGAGGTGGGCGGCCAGTACGCCCTGGGCATCATCCAGGCGCG    457
             ||||| ||||||| |||||||||||||||||||| ||| |||||||||||| ||| 
Qy   241 GCCCCCTGGTGACCATCAGGATCGGCGGCCAGTACGCCCTGGGCATCATCCAGGCCG     300

Qy   458 CCGACGACACCGTGCTGGAGGATGAGCCTCGCCGGCAAGTGAAGCCCAAGATGATCG    517
             |||| ||| ||||| |||| |||| |||| ||||||||||||||| ||||||||||| 
Db   301 CCGACGACACCGTGCTGGAGGATGAACCTGGCCAGCAAGTGGAGCCCAAGATGATCG     360

Db   518 GCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCG    577
             |||||||||||||||||||||||| |||| || ||||| |||| |||||||| ||||| 
Qy   361 GCGGGATCGGCGGCTTCATCAAGGTGCGCCAGTACGAGCAGATCCCCGTGGGAGATCTGCG    420

Qy   578 GCAAGAAGGCCATCGGCGTGCTGATCGGCCCCGCACCCCGTGAACATCATCGGCCCGCA    637
             |||| ||||||||||||||||| ||||||||| |||| |||||||||||||||||||| 
Db   421 GCCACAAGGCCATCGGCGTGCTGATCGGGCAGTGCGCCCCATCAGCCCCGTGAACATCATCGGCCGCA     480

Db   638 ACATGCTGACCCAGCTGGCTGCACCCTGAACTTCCCAAGGTGAAGCAGGTGAAGCAGCTGG     697
             |||||||||||||| ||||| ||| ||||||||||| |||| | ||| |||||||||| 
Qy   481 ACCTGCTGACCCAGATCCGGCTGCACCCTGAACTTCCCAAGGTGCAGTCAGCTGGACCGTGC    540

Qy   698 CCGTGAAGTGAAGCGCCCGCATGGACGGCGGGGATGACGCGGCCCTGACCGAGG    757
             ||||||||||||||||||||||||| |||| ||||| ||||||||||||||||| 
Db   541 CCGTGAAGTGAAGCGCCCGCATGGACGGCGGGGATGACGCGGCCCTGACCGAGG     600

Db   758 AGAAGATCAAGGCCCTGACCGCCATCTGCGAGAGATGGAGAGGCAAGATCACCA     817
             |||||||||||||| || |  |  |  || ||||||||||||||||||||||||  
Qy   601 AGAAGATCAAGGCCCTGTTGGAGATCTCACCGAGGATGGAGAGGCAAGATCAGCA    660

Qy   818 AGATCGGCCCGGAGAACCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGACAGCA    877
             |||| ||||| ||||||||||| |||||||||||||||||||||||||||||||||| 
Db   661 AGATCGGCCCGGAGAACCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGACAGCA     720

Db   878 CCAAGTGGCGCAAGCTGGTGGATTTCCGCGAGTGAACAAGCCGCACCCAGGACTTCTGGG    937
             ||||||||||||||| || |||||||||| || ||||| ||| |||| |||||||||||| 
Qy   721 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGTGAACAAGCCGCACCCAGGACTTCTGGG     780

Qy   938 AGGTGCAGCTGGGCATCCCCACCCCCGCCTGAAGAGAGAGAGCGTGACCGTGC    997
             |||||||||||||||| |||||||||||| |||||||||||||||||||||||| 
Db   781 AGGTGCAGCTGGGCATCCCCACCCCCGCCTGGCTGAAGAGAGCGTGACCGTGC     840

Db   998 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCTGCCCCGGACAAGGACTTCCCCAAGTACACCG    1057
             ||||||||||||| ||||||||||||||| |||||||| |||||||| ||||||||||||||| 
Qy   841 TGGACGTGGGCGACGCCTACTTCAGCGTGCCCTTCCCCGGACAAGGACTTCCCCAAGTACACCG     900

Qy  1058 CCTTCACCATCCCCAGCATCAACAACGAGACCCCGGCATCCCCGGCATCCGCTACCAGTACAACGTGC    1117
             |||||||||||||||||||||||||||||| ||||||||||||||||| ||||||||||||||||||| 
Db   901 CCTTCACCATCCCCAGCATCAACAACGAGACCCCGGCATCCGCTACCAGTACAACGTGC     960

Db  1118 TGCCCCAGGGCTGAAGGGCAGCCCCAGCATCTTCCAGACCAGCAGCATGACCAAGATCCTGG    1177
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Qy   961 TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGACCAGCAGCATGACCAAGATCCTGG    1020

Qy  1178 AGCCCTTCCGCGCCCAGAATCGTGATCTACCAGTACAACGATGACAGTACAACGTACG    1237
             |||||||||||||||||||| ||||| ||||||||||      |||||||||||||||
Db  1021 AGCCCTTCCGCAAGCAGAACCCCGACATCGTGATCTACCA------GGCCCCTGTACG    1074
```

RESULT 3
US-09-475-515-83

; Sequence 83, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 2300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(-).protmod.RTopt.YMWM
US-09-475-515-83

Query Match 80.9%; Score 1998.4; DB 4; Length 2300;
Best Local Similarity 92.9%; Pred. No. 0;
Matches 2147; Mismatches 141; Indels 24; Gaps 4;

```
Db   2149 ACAAGCTGGTGAGCGGCCGCCGGCATCCGCAAGGTGCTGTTCCTGAACGGCATCGATGGCGGCA 2208
          |||||||||||||

Qy   2378 TCGTGATCTACCAGTACATGGACGACCTGTGGGTGGGCAGCGGCGGCCCTAGGATCGATT 2437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2209 TCGTGATCTACCAGTACATGGACGACCTGTGGGTGGGCAGCGGCGGCCCTAGGATCGATT 2268

Qy   2438 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2469
          |||||||||||||||||||||||||||||||
Db   2269 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2300
```

RESULT 4
US-09-475-515-6
; Sequence 6, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
;   APPLICANT: BARNETT, Susan
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: SRIVASTAVA, Indresh
;   APPLICANT: LIAN, Ying
;   APPLICANT: HARTOG, Karin
;   APPLICANT: LIU, Hong
;   APPLICANT: GREER, Catherine
;   APPLICANT: SELBY, Mark
;   APPLICANT: WALKER, Christopher
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;   TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;   FILE REFERENCE: 1621.002
;   CURRENT APPLICATION NUMBER: US/09/475,515A
;   CURRENT FILING DATE: 1999-12-30
;   NUMBER OF SEQ ID NOS: 90
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 4319
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: HIV-Gag-polymerase
US-09-475-515-6

  Query Match   79.8%; Score 1969.4; DB 4; Length 4319;
  Best Local Similarity   90.1%;
  Matches 2158; Conservative   0; Mismatches 221; Indels   16; Gaps 4;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCAGC--GCCAACATCCTGATGCAGCGCCAGCAACT 70
          |||||||||| ||||||||||| || |||| ||| ||| ||||||||||||| || ||||
Db   1100 TGGCCGAGGCGGATGAGCCAGGTGACGAACCCGGACCATCATGATGCAGCGGCGGCAACT 1159

Qy   71 TCAAGGGCCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
          || ||| |||||||||| || |||||||||||||||||||||||||||||||||| |||||
Db   1160 TCCGCAACCAGCGCAGCGGCGGCATGAGTGCTTCAACTGCGGCAAGGAGGGCCACCACCGCCA 1219

Qy   131 GCAACTGCGCGCCCCCCCGCCGCCGCAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGA 190
          | ||||| ||||| |||||| ||||||| | ||||| | ||| |||||||||||||||||
Db   1220 GGAACTGCCGCGGCCGCCGCCGCCGCAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAAA 1279

Qy   191 TGAAGGACTGCACCGAGCGCCAAGCTTCTTCCGCGAGGACCTGGCCTTCCCCCAGG 250
          |||| |||| || ||| || | ||| |||| |||||| ||||| ||||||||
Db   1280 TGAAAGATTGCACTGAGCAGGCAGGAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCACCAAG 1338

Qy   251 GCAAGGCCGCGAGTTCCCCGAGCGAGCGCAGAACCGCGCCACAGCCCGCGCGAGC 310
          || ||||| |||||| || || |||||||||| | |||| || ||||| |||||
Db   1339 GGAAGGCCCAGGGAATTTCTTCAGACGCAGCAGAGCCAACCAGCCCACCCAGCGAGAGC 1398

Qy   311 TGCAGGTGCGCGGCG------ACAACCCCGCCAGCGAGCGCCAGGGCA 364
          |||| | |||| ||| ||| ||| | ||| |||| |||||||||
Db   1399 TTCAGGTTGGGGAGGAGAAAACAACTCCCTCACCGTCGTCGATAGACAAGGAA 1458

Qy   365 -----CCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTCACGTGTAAGGA 418
               |||| |||| |||| |||| |||| ||| | ||||| |||| | || ||| |||||
Db   1459 CTGTATCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTGGTCACAGTGAAGGA 1518
```

```
Qy   1238 TGGGCAGCGACCTGAGATCGGCGCCGCCGCCAGCCGCCACCGCGCCAGATCGAGGAGTGCGCCAAGCACC 1297
          |||||||||||||||||||||||| ||||||||| |||  ||| ||| | |||| |||||||||
Db   1075 TGGGCAGCGACCTGAGATCGGCGGCCGCCGCCAGCGCCACCGCGCCAGATCGAGGAGTGCGCCAGCACC 1134

Qy   1298 TGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGT 1357
          |||||||| |||| |||||||||||| ||||||||||||| ||||||||| ||||| |||
Db   1135 TGCTGCGCTGGGGCTTCACCACCCCGGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGC 1194

Qy   1358 GGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGTCCTGCCCGAGA 1417
          |||| |||||| | | |||  ||||| ||| ||||||||||   ||||| ||||||||||
Db   1195 CCAT------CGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGTCCTGCCCGAGA 1248

Qy   1418 AGGAGAGCTGGACCGTGACCGACATCCAGAAGTTGGTGGCAAGCTGAACTGGGCCAGCC 1477
          || |||||||||||||| ||||||  |||| ||| ||||||||||| ||||||||||||
Db   1249 AGGACAGCTGGACCGTGACCGACATCCAGAAGTTGGTGGCAAGCTGGAACTGGGCCAGCC 1308

Qy   1478 AGATCTACCCGGCCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCACCAAGGCCC 1537
          ||||||||||||||||||||||| || ||||||| ||||| ||| ||| ||| |||||||
Db   1309 AGATCTACCCGGCCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCACCAAGGCCC 1368

Qy   1538 TGACCGACATCGTGCCCCTGACCGAGGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGA 1597
          ||||||||||||||| || || |||||||| | |||||||||| ||||||||||||||||
Db   1369 TGACCGAGGTGATCCCCCTGACCGAGGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGA 1428

Qy   1598 TCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGA 1657
          ||||| ||||||||| |||| ||| ||||| ||| ||||||||||||||||||||||||||
Db   1429 TCCTGAAGGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGA 1488

Qy   1658 TCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACC 1717
          ||||||||| ||||| ||||| ||| |||||||||||||||||||||||||||||||||||
Db   1489 TCCAGAAGCAGGGGCAGGGGCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACC 1548

Qy   1718 TGAAGACCGGCAAGTACGCCAAGATCCGCACCACCACCGAGGCAGTGA 1777
          || |||||||||||||| |||| | |||||| ||| ||| |||||||||
Db   1549 TGAAGACCGGCAAGTACGCCCGCATGCGCGGCCCACCACCGAGGCAGTGA 1608

Qy   1778 CCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGCCAAGACCCCCAAGT 1837
          |||| | |||||||||| |  |||  |||||||||||||| ||| |||||||||||||||
Db   1609 CCGAGGCCGTGCAGAAGGTGAGCATCAGAGCATCGTGGTCTGGGCCAAGACCCCCAGT 1668

Qy   1838 TCCGCCTGCCCATCCAGAAGGAGACCTGGGAGGCCTGGTGGACCGAGTACTGGCAGGCCA 1897
          |||||||| ||||||||||||||||| |||||| ||| |||||||||||||||||||||||
Db   1669 TCAAGCTGCCCATCCAGAAGGAGACCTGGGAGGCCTGGTGGACCGAGTACTGGCAGGCCA 1728

Qy   1898 CCTGGATCCCGAGTGGGGAGTTCGTGAACACCCCCCCTGGTGAAGCTGTGGTACCAGC 1957
          ||||||||||| |||||||||| ||||| ||| ||||| || |||| |||||||||||
Db   1729 CCTGGATCCCCGAGTGGGGAGTTCGTGAACACCCCCCCTGGTGAAGCTGTGGTACCAGC 1788

Qy   1958 TGGAGAAGGAGCCCATCGTGGGCGCCGAGACGCCGCCAACCGCG 2017
          ||||||||||| ||||  |||||| |||| |||| || ||||||
Db   1789 TGGAGAAGGAGCCCATCGTGGGCGCCGAGACGTTCTACGTGGACGGCGCCGCCAACCGCG 1848

Qy   2018 AGACCAAGATCGGCAAGGCCGGCTACGTGACCAACCGGGGCCGGCAGAAGGTCGTGAGCC 2077
          | |||||||| || |||||||||||| ||| ||| ||||| ||| ||| | || |||| |
Db   1849 AGACCAAGATCGGCAAGGCCGGCTACGTGACCAACCGGGGCCGGCAGAAGGTCGTGACCA 1908

Qy   2078 TGACCGACACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACA 2137
          ||||| | | ||| | || |||||  |||||  ||| || | || |||||||||||| |
Db   1909 TGACCGACACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACA 1968

Qy   2138 GCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCGGGGCAGC 2197
          ||||||| |  ||||||  | ||||||| |||||||||||| ||| ||||| |||||||||
Db   1969 GCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCC 2028

Qy   2198 AGCCCGACGTGCAGGGCGAGCGCCACACCGGCAGCGCGCCACAGAGATCGAGGAGGA 2257
          |||  |||| | | ||| |||||||| |||| |||||||||||| ||||||||||||
Db   2029 AGCCCGACGTGCAGGGCGAGCGCCACACCGGCAGCGCGCCACAGAGAGATCGAGGAGGA 2088

Qy   2258 AGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCG 2317
          ||| |||| || |||||||||||| | |||||||||||||||||||||| ||||||| |
Db   2089 AGAAGGTGTACCTGGGTGCCCTGGTGGTCCGGCCCAGCAGGCCATCGGCGGCAACGAGCAGCAGTGG 2148

Qy   2318 ACAAGCTGGTGAGCGGCCGCCGGCATCCGCAAGGTGCTGTTCCTGAACGGCATCGATGGCGGCA 2377
          |||| || |||||||||| ||||| |||  |||||||||||| ||| ||||||||  | || | 
Db   2318 ACAAGCTGGTGAGCGGCCGCCGGCATCCGCAAGGTGCTGTTCCTGAACGGCATCGATGGCGGCA 2377
```

```
Qy  1499  TGCGCCAGCTGTGCAAGCTGCTGCGGGGGCCAAGGCCCTGACCGACATCGTGCCCCTGA  1558
          ||||||||||||||||||   |||||||||  |||||||||||||||| |||||||||||
Db  2599  TGAAGCAGCTGTGCAAGCTGCTGCGGCGGCCACCAAGGCCCTGACCGAGTGATCCCCTGA  2658

Qy  1559  CCGAGGAGGCCGAGCTGAGCTGGCCGGAGATCCTGCGGGAACCGCGAGACCCGTGCACG  1618
          ||||||||||||||||   |||  |||||||||||||||||||||||||||||||||||
Db  2659  CCGAGGAGGCCGAGCTGGAGCTGGCCGGAGATCCTGAAGGAGCCCCGTGCACG        2718

Qy  1619  GCGTGTACTACGACCCCGAGCAAGGACCTTCGGCAACGCCATCAGAAGCAGGGCCACGACC  1678
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  2719  AGGTGTACTACGACCCCAGCAAGGACCTTGGTGCCGAGATCCAGAAGCAGGGCCCAGGGCC  2778

Qy  1679  AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCA  1738
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  2779  AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCC  2838

Qy  1739  AGATGCGCACCGCCCACACACCAACAGCTGAAGCAGTGACCGAGCAGCCGCCGTGCAGAAGATCG  1798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2839  GCATGCGCGGCGCGCCCACACCAACAGCTGAAGCAGTGACCGAGCAGCCGCCGTGCAGAAGGTGA  2898

Qy  1799  CCATGAGAGCAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCTGCCATCCCGAGTGGGAGG  1858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2899  GCACCGGAGAGCAGCATCGTGATCTGGGTGGATCAAGATCAAGTTCAAGCTGCCATCCCGAGTGGGAGT  2958

Qy  1859  AGACCTGGAGACCTGGTGGACCACTACTGGCAGGCGCGCGCCCGCCACCTGATCCCGAGTGGGAGT  1918
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2959  GCACCTGGAGAGCGCCTTGGTGATGGAGTACTGGCAGGCGCGCGCCCGCCACCTGGATCGTGG  3018

Qy  1919  TCGTGAACACACCCCCCCCGTGTGAAGCTGTGGTACCAGCGCGGAGAAGGAGCCCCATCATCG  1978
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3019  TCGTGAACACACCCCCCCCGTGTGAAGCTGTGGTACCAGCTGGAAGGAGCCCATCCGTGG  3078

Qy  1979  GCGCCCAGACCTTCACGTGACGGCCGGACCAAGAATCGGCAAGGCCG  2038
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3079  GCGCCCAGACCTTCACGTGACGGCCGGACCAAGCTGGGCAAGGCCG  3138

Qy  2039  GCTACGTGACCGACCGGCGGCCGGCAGAAGATCGTGAGCCTGACCGAGCACCACCAGA  2098
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3139  GCTACGTGACCGACCGGCGGCCGGCATCGCCGGAGGTGTGAGCATCGCCGACACCACCAGA  3198

Qy  2099  AGACCCGAGCTGCAGGCAGCTGCCATCCAGCTGCCCTTCACCATCCCAGCATCG  2158
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3199  AGACCCGAGCTGCAGGCGTGCCATCCAGCTGCCCTTCACCATCCCAGCATCA  3258

Qy  2159  TGACCGACAGCCAGTACGCCCTGGCATCATCCAGGCCGACAAGAGAGGGAGCG  2218
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3259  TGACCGACAGCCAGTACGCCCTGGCATCATCCAGCGCCCCGACAACAGAGCGAGAGCG  3318

Qy  2219  AGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGAAGTGTACCTGGAGCTGGG  2278
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3319  AGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGAAGTGTACCTGCGCCTGGG  3378

Qy  2279  TGCCCGCCACAAGGGCATCGGCCAGCTGGCGGAGTGGGCAGGGCCAGGAGGTTGTGAGCAAGGGCA  2338
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3379  TGCCCCGCCACAAGGGCATCGGCCAGCTGGTGGCAGGGCCAGGAGGTGGCAGGGCAGGCGCCGGCA  3438

Qy  2339  TCCGCAAGGTGCTGTTCCTGACGCCATCGTGATCGTACCAGTA  2393
          |||||||||||||||||||||||||||||||||||||||||||
Db  3439  TCCGCAAGGTGCTGTTCCTGAACGCGATCCTGACGCCGAGGAGCACGAGAA  3493
```

RESULT 5
US-09-475-515-80
; Sequence 80, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine

```
Qy  419   TGGGGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCCTGCTGGAGG  478
          ||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db  1519  TCGGCGGCCAGCTCAAGGAGGAGGCGCTGCTGCGCCGACACCGGCACCGGTGCTGGAGG  1578

Qy  479   AGATGAGCCTGCCGGCAAGTGGAAGCCCAAGATCGGCGGCCATCGGCGGCTTCATCA  538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1579  AGATGAACCTGCCCGGCGAAGTGGAAGCCCAAGATGATCGGCGGGATCGGGGGCTTCATCA  1638

Qy  539   AGTGCGCGCAGTACGACCAGATCTGATCGAGATCTCGCGGCAAGAGAGGCCATCGGCACCG  598
          ||||||||||||||||   ||||||||||||||||||||||||||||
Db  1639  AGTGGCGGCCAGTACGACCAGATCCCCGTGGAGATCTCGCGGGATCTGCGGCCACAAGGGCCATCGGCCACCG  1698

Qy  599   TGCTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCT  658
          ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db  1699  TGCTGGTGGGCCCCACCCCCGTGAACATCATCGGCCGCAACATCCTGCTGACCCAGCTGGGCT  1758

Qy  659   GCACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCA  718
          ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db  1759  GCACCCTGAACTTCCCCATCGAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGGA  1818

Qy  719   TGGACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG  778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1819  TGAAGGGCCCCAAGGTCAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGTGTGG  1878

Qy  779   CCATCTGCGAGGAGATGGAGAAGGACCAAGATCACCAAGACCAAGATCAAGGCCCTGTGTGG  838
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1879  AGATCTGCACCGGAGATGGAGAAGGACCAAGATCAGCCAAGATCAAGGCCCTGTGTGG  1938

Qy  839   ACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG  898
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1939  ACAACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG  1998

Qy  899   ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC  958
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1999  ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC  2056

Qy  959   ACCCCGCGGCCTGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACT  1018
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2059  ACCCCGCGGCCTGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACT  2116

Qy  1019  TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACACCGCCTTCACCATCCCAGCATCA  1078
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2119  TCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACACCGCCTTCACCATCCCAGCATCA  2178

Qy  1079  ACAACGAGACCCCCGGCATCCGGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA  1138
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2179  ACAACGAGACCCCCGGCATCCGGTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCA  2238

Qy  1139  GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACC  1198
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2239  GCCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCAAGCAGAACC  2298

Qy  1199  CCGAGATGCTGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCG  1258
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2299  CCGACATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCG  2358

Qy  1259  GCCAGCACCGCGCCAAGATCGAGGAGCTGCGCGAACACCTGCTGCGCTGGGGCTTCACCA  1318
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2359  GCCAGCACCGCGCCACCAAGATCGAGGAGCTGCGCGAACACCTGCTGCGCTGGGGCTTCACCA  2418

Qy  1319  CCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACC  1378
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2419  CCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACC  2478

Qy  1379  CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGTGACCGTGAACG  1438
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2479  CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACAGCTGGACCGTGAACG  2538

Qy  1439  ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGATCTACCCCGGCATCAAGG  1498
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2539  ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCCAGATCTACGCCGGCATCAAGG  2598
```

```
;  APPLICANT: SELBY, Mark
;  APPLICANT: WALKER, Christopher
;  TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
;  TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
;  FILE REFERENCE: 1621.002
;  CURRENT APPLICATION NUMBER: US/09/475,515A
;  CURRENT FILING DATE: 1999-12-30
;  NUMBER OF SEQ ID NOS: 90
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 80
;  LENGTH: 2305
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: FS(+).proinact.RTopt.YM
US-09-475-515-80

Query Match         75.3%;  Score 1858;  DB 4;  Length 2305;
Best Local Similarity 89.0%;  Pred. No. 1.2e-298;
Matches 2058;  Conservative 0;  Mismatches 235;  Indels 19;  Gaps 4;
```

```
Qy    170  GCGGCAAGGAGGGCACCAGATGAAGAGGACTGACCAGGAGGCCAACTTCTTCCGCG       229
                ||| || || ||||| |  |||||| || ||  ||  ||||||  ||   ||| |||
Db      1  GCGGCCGCGAAGGACACCAAATGAAAGATTGCACTGAGAGACAGGCTAATTTTTT-AGGGG  59

Qy    230  AGGACCTGGCCTTCCCCCAGGGCAAGGGCCGCGAGTTCCCCAGCGAGGACCGCGGCCA     289
               ||| | ||||||  | ||||| ||||| ||  | ||   ||||||||  | | ||||||
Db     60  AAGATCTGGCCTTCTACAAGGGAAGGCCAGGGAATTTTCTTCAGAGCGAGACCAGAGCCA    119

Qy    290  ACAGCCCCACCAGCCGCGAGCTGCAGGTCGCGGCG-----ACAACCCCGCCAGCGAGG     343
               ||||||||||||| | |||||| |||| | |||| ||   |||||   | || | |||| |
Db    120  ACAGCCCCACCAGCAGAGAGCTTCAGGTTTGGGGAGGAGAAAACAACTCCCTCTCAGAAG    179

Qy    344  CCGGCGCCGAGCGCCAGGGCA------CCCTGAACTTCCCCAGTCACCCTGCGCAGC     397
               || | ||| ||  |  ||| |     ||| |||||||||  || ||  |||   |||||
Db    180  CAGGAGCCGCATAGACAAGAACTGTATCTTTAACTTCCCTCAGATCACTCTTTTGGCAAC    239

Qy    398  GCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCG     457
               || || || |  | || ||||  |  || ||||||||||||||| ||| || | |||| ||
Db    240  GACCCCTGCTGCACAATAAGGATCGGGGCGTTCATCAAGGAGGAGCCGCTGCTGATACAGGAG   299

Qy    458  CCGACGACACCGTGCTGGAGGATGAGCTGCAGGTCGGCCAAGTGAAGCCAAGATGATCG     517
               || | ||| ||||| | ||||  ||||| ||||  || ||  |||| ||  ||||| | |
Db    300  CAGATGATCAGTATTAGAGAAGAAATGAATTTGCCAGGAAAATGAAACCAACAATGATAG    359

Qy    518  CCGGCATCGGCGGCTTCATCAAGGTGCGTGGGAGGATGAGCCTGCCGCAGTACGACCAGGTG       577
               || | || |||  ||||||||||  | || || | || | || | |  || | | |||| |
Db    360  GGGGGATCGGCGGCTTCATCAAGGATGTGGGAGGCAGTACGACCAGGTACGACCAGATAGTG  419

Qy    578  GCAAGAAGGCCATCGGCGGCTCATCGATGGCCACCCGCACCCCCGTGAACATCATCGGCGCA    637
               ||  |||||| | || |||| | || | | | ||  | | ||  ||||| || || |||| |
Db    420  GACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAA    479

Qy    638  ACATGCTGACCGAGCTGGGTGCACCTGAACTTCCCATCAGCCCATCAGCGAGACCGTGC     697
               || | || ||| |  || || || |||  |||||||| |||| |||  || |||| ||||||
Db    480  ATCTGTTGACCCAGATCGGCGGCTGCACCTTGAACTTCCCCATCAGCCCATTGAGACGGGTGC  539

Qy    698  CCGTGAAGCTGAAGCCCGGCATGGACGGCCCGAAGGTGAAGCAGTGGCCCCTGACCGAGG    757
               |||| |||| || |||||||||||  |||||| || || || | |  |||||| |||||||
Db    540  CCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTCAAGCAATGGCCATTGACCGAGG    599

Qy    758  AGAAGATCAAGGCCCTGACCGAGATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCA    817
               || ||||||| |||||| ||||| |  | || || |  || || ||  | || | |  |||
Db    600  AGAAGATCAAGGCCCTGGTTGAGATCTGCACCGAGATGGAGAAGGAGGGCAAGATCAGCA    659

Qy    818  CCGATCGGCCCGAGAACCCTACAACACCCGTGTTCGCCATCAAGAAGAAGGACAGCACA    877
               ||||| | ||| |||||||| | |  | ||  || || || |||||||| | | |||| |
Db    660  AGATCGGCCCCGAGAACCCCTACAACACCCCGGTGTTCGCCATCAAGAAGAAGGACAGCACA    719

Qy    878  CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGAGCACCCAGGACTTCTGGG    937
               |||||||| | |||| || |||||| | |||||||| ||||||||||| |||||| | |||
Db    720  CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGAGCACCCAGGACTTCTGGG    779
```

```
Qy    938  AGGTCAGCTGGCATCCCCACCCCCACCCGCGCCTGAAGAGAGAAGAGAGCTGACCGTGC       997
               ||||||||||||| ||||||||||||||| |||| | |||||||||| |||| |||||||
Db    780  AGGTCAGCTGGCATCCCCACCCCCACCCGCCTGAAGAGAAGAGAGCTGACCGTGC         839

Qy    998  TGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGCCCCTGACAACGGAGGACTACACCG    1057
               ||||||||||| ||| |||||||||||||||||||| |||||||| | ||||||||||||
Db    840  TGGACGTGGGGGACGCGCTACTTCAGCGTGCCCCTGCCCCTGACAAGGAGGACTACACCG    899

Qy   1058  CCTTCACCATCCCAGCATCAACAACGAGAGCCCCGGCATCCGCTACCAGTACAACGTGC    1117
               ||||||| || |||||| | | ||| ||| |||| |||||||| | |||| | |||||||
Db    900  CCTTCACCATCCCAGCATCAACAACGAGAGCCCCGGCATCCGCTACCAGTACAACGTGC    959

Qy   1118  TGCCCCAGGGCTGGAAGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG    1177
               ||| |||||||| || |||  |||||||||||||||| || | |  ||||||  ||||||
Db    960  TGCCCCAGGGCTGGAAGGGCAGCCCCATCTTCCAGAGCAGCATGACCAAGATCCTGG    1019

Qy   1178  AGCCCTTCCGGGCGCAACCCGGCGATCGTGATCTACCAGTACATGGACGACCTGTACG    1237
               |||||||||| |  ||||| |||  | || ||||||||| |||||||||||||||||||
Db   1020  AGCCCTTCCGCAAGCAGAACCCGACATCGTGATCTACCA-----GGCCCCGTGTACG    1073

Qy   1238  TGGGCAGCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGCCAGCACC    1297
               ||||||||||| || ||||||||||||  |||||||||||||||| ||||||  |||||||
Db   1074  TGGGCAGCGACCTGGAGATCGGCCAGCACCGCACCAAGATCGAGGAGCTGCGCCAGCACC    1133

Qy   1298  TGCTGCGCTGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGT    1357
               ||||||||||||| ||||||||||||||||||||| ||||| ||||||||||||||||||
Db   1134  TGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGT    1193

Qy   1358  GGATGGGCTACGAGCTGCACCCGAGCAAGTGGACGGTGCAGCCGATCGTGCTGCCCGAGA    1417
               |||||||||| ||| || || ||| ||||| ||||||||||||||||||||||||||||||
Db   1194  GGATGGGCTACGAGCTGCACCCGGACAAGTGGACGGTGCAGCCGATCATGCTGCCCGAGA    1253

Qy   1418  AGGAGAGCTGACCGTGAACGACATCAGAGAGCTGGTGGGCAAGCTGAACTGGGCCAGCC    1477
               ||||| ||||||||||||||||||||| ||||| || |||||||| || ||||||||||||
Db   1254  AGGACAGCTGACCGTGAACCGGACGATCATCCAGAGCTGGTGGGCAAGTGAACTGGGCCAGCC    1313

Qy   1478  AGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGGCGCCAAGGCCC    1537
               ||||||| |||| |||||||| |  || || | |||||| |||||||||| || ||||||||
Db   1314  AGATCTACCCGCCGATCTACAAGGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCACCAAGGCCC    1373

Qy   1538  TGACCGACATCGTGCCCCTGACCGAGGCGCAAGTGAAGGCCAAGATGATCG    1597
               ||||||| |||||||||||| | || || |||| |||| ||||| | | | |
Db   1374  TGACCGACGTGATCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGGGAGA    1433

Qy   1598  TCCTGCCGGAGCCGTGCTACTGGACCGCGTGTTACTACGACCCCAGCAAGGACACCGTGGCGGCCGAGA    1657
               |||||| |||| ||| ||| | | || ||||| | ||||||||| || ||| |||||||||| ||||||
Db   1434  TCCTGAAGGAGCCGTGCTACTGGACCGCGTGTTCGACGACCCCAGCAAGGACCTGGTTGCTGGTCCGCCGAGA    1493

Qy   1658  TCCAGAAGCAGGGCCAGGGCCAGTGGACCTTCTACCAGATCTACCAGGAGCCCTTCAAGAACC    1717
               |||||| ||||| |||| || |||||||||||  | | | ||  |||||||||||||||||||
Db   1494  TCCAGAAGCAGGGCCAGTGGACCTTCTACCAGATCTACCAGGAGCCCTTCAAGAACC    1553

Qy   1718  TGAAGACCGGCAAGTACGCCAAGATGCGCAGCGCGCACACCAACGACGTGAAGCAGCTGA    1777
               ||||| |||||||| || || |||| | | |||||| |||  ||| | | ||||||||||
Db   1554  TGAAGACCGGCAAGTACGCCAAGATGCGCGGCGCCCACACCAACGACGTGAAGCAGCTGA    1613

Qy   1778  CCGAGGCCGTGCAGAAGATCGCCACCGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGT    1837
               ||||| | ||||| |||| ||| | ||||| ||  || | |||||||||| ||||||||||
Db   1614  CCGAGGCCGTGCAGAAGATCGCCACCGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGT    1673

Qy   1838  TCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAGTACTGGCAGGCCA    1897
               |||| || ||| ||||||||||||||  | || |||  |||| ||  | ||||||||| ||
Db   1674  TCAAGCTGCCCATCCAGAAGGAGACCTGGGAGACGTGGTGGACCGAGTACTGGCAGGCCA    1733

Qy   1898  CCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGGTGTGAAGCTGTGTGTACCAGC    1957
               || || |||||||||| ||||| ||||||||||||||||||||| | | ||||||||||||||
Db   1734  CCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGGTGTGTGTGAAGCTGTGTGTACCAGC    1793

Qy   1958  TGGAGAGGAGCCCATCATGGCGCGGCCCGAGACCTTCTACGTGACGGCGCGCCAACCGCG    2017
               |||||||||||||||||| |||| ||||||||||||||| |||| | | |  | ||||||
Db   1794  TGGAGAAGGAGCCCATCGTGGGCGCGGCCCGAGACCTTCTACGTGACGGCGCGCCAACCGCG    1853
```

```
Qy      60 AAGATCTGGCTTCCTACAGAGGAAGGCCAGGAGGGAATTTTCTTCAGAGCGAGCCAGAGCCA 119

Db     290 ACAGCCCCACCAGCCGCGAGCTGACGTGCGCGCG----ACAACCCCAGCGAGG 343
Db     120 ACAGCCCCACCAGCCGCGAGCTGACGTGCGCGCG----ACAACCCCAGCGAGG 179

Qy     344 CCGGCGCGAGGCGCAGGGCA------CCCTGAACTTCCCCAGATCACCCTGGCGCAGC 397
Db     180 CAGGAGCCGGATAGACAAGGAACTGTATCTTTAACTTCCCTCAGATCACTCTTTGGCAAC 239

Qy     398 GCCCCCTGTGAGCATCAAGGTGGGCGGCCAGTACGCCCTGGACACCGGCG 457
Db     240 GACCCCTGTCACAATAAGGATCGGGGGCAACTCAAGGAAGCGCTGCTGTCGATACAGGAG 299

Qy     458 CCGACGACACCGTGCTGGTGAGGAGGAGGATGAGCGTGCCGCCGAAGTGGAAGCCCAA 517
Db     300 CAGATGATACAGTATTAGAAGAAATGAATTGCCAGGAAAATGAAAACAAAAATGATAG 359

Qy     518 GCGGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCTGCGG 577
Db     360 GGGGGATCGGGGGCTTCATCAAGGTGGAGGCAGTACGACCAGATACCTGTGAAATCTGTG 419

Qy     578 GCAAGAAGGCCATCGGCACCGTGATCGCCACCCGTGAACATCATCGGCCGCA 637
Db     420 GACATAAAGCTATAGGTACAGTATTAGTAGACTTCAACAATAATTGGAAGAA 479

Qy     638 TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGGCAGCCCATCAGCCCATCAGCCCTCCTATTGGGGATCGATT 697
Db     480 ATCTGTTGACCCAGATCGGCTGCACTGGAACTTGAACTTCCCCAAGGTCAGTCAATGGCCCATTGACCGAGG 539

Qy     698 CCGTGAAGTGAAGCGGGCATGACGGCCCAAGGTGAAGCAGTGGCCCCTGACCGAGG 757
Db     540 CCGTGAAGTTGAAGCGGGATGGACGCGGGGCCCAAGGTCAAGCAATGGCCCATTGACCGAGG 599

Qy     758 AGAAGATCAAGGCCCTGACCGACGCCATCTGCGAGGAGGATGAGGAGGGCCAAGATCACCA 817
Db     600 AGAAGATCAAGGCCCTGGTGTGAGATCTGCACGAGATCTGACCGAGATGGAGGAGGGCCAAGATCAGCA 659

Qy     818 AGATCGGCCCCGAGAACCCTACAACACCCGTGTTCGCCATCGAACAGCAGA 877
Db     660 AGATCGGCCCCGAGAACCCTACAACACCCGTGTTCGCCATCGAACAGCAGA 719

Qy     878 CCAAGTGGCCGCAAGTGGTGTGACTTCCGCAGCTGAACAGGGCCACCCAGGACTTCTGGG 937
Db     720 CCAAGTGGCGCAAGTGGTGGACTTCCGCAGCTGAACAGCGACCCAGGACTTCTGGG 779

Qy     938 AGGTGCAGCTGGCCATCCCCGGCCTGAAGCAGCGCCTGAAGAGAGAGGGTGACCGTGC 997
Db     780 AGGTGCAGCTGGCATCCCCACCCCGGCCCTGAAGCAGCGCCTGAAGAGAGAGGGTGACCGTGC 839

Qy     998 TGGACGTGGGCGGACGGCCTACTTCAGCGTGCCCTGCGACGTACACCG 1057
Db     840 TGGACGTGGGCACGCCTACTTCAGCGTGCCCTGCGACAAGGTACACCG 899

Qy    1058 CCTTCACCATCCCCAGCATCAACAACGAGCCCCGGCATCCGGCTACCAGTACACAACGTGC 1117
Db     900 CCTTCACCATCCCCAGCATCAACAACGAGCCCCGGCATCCGGCTACCAGTACACAACGTGC 959

Qy    1118 TGCCCAGGGCTGGAAGGCAGCCCAGCATCTTCAGAGCAGCAGCAAGATCCTGG 1177
Db     960 TGCCCCAGGGCTGGAAGGGCAGCCCCATCTTCAGAGCAGCAGCAAGATCCTGG 1019

Qy    1178 AGCCCTTCCGCGCCGCAACCCGAGATCGTGATCTACCAGTACGATCATCTACCA 1237
Db    1020 AGCCCTTCCGCGCCGCAGCAGAACCCGACATCGTGATCTACCA------GGCCCCCGTGTACG 1073

Qy    1238 TGGGCAGCAGACCTGGAGATCGGCCAGACGCCCCAAGATGGAGAGCTGCGCAAGCACC 1297
Db    1074 TGGGCAGCAGGACCTGGAGATCGGCAGCACCAGAAGATGAGAGAGCTGCGCAAGCACC 1133

Qy    1298 TGCTGGGCTCGGGCTTCACCACCCCAGAGTTCCCCAGCAGCGAGCAGAAACCGCGCCA 1357
Db    1134 TGCTGGGCTCGGGCTTCACCACCCCAGAGAGGAGGCCAGAAGGACCAGAGGAGCCCCCCTTCTGC 1193
```

```
Qy    2018 AGACCAAGATCGGCAAGGCCGGCTACTGACCGACGCGGGGGCCGGCAGAAGATCGTGAGCC 2077
Db    1854 AGACCAAGCTGGGCAAGGCCGGCTACTGACCGACGCGGGGGCCGGCAGAAGGTGTGAGCA 1913

Qy    2078 TGACCGAGACCACCAACCAGCAGAGACCGAGCTGCAGGCCATCCAGCTGCGGCCCTGCAGGACA 2137
Db    1914 TCGCCGACACCACCAACCAGAGAGACCGAGCTGCAGGCCATCCAGCTGCGGCCCTGCAGGACA 1973

Qy    2138 GCGGCAGCAGCGAGGTGAACATCTGACGAGCCAGTACGCCCTGGCATCATCCAGGCCC 2197
Db    1974 GCGGCCCTGGAGGTGAACATCTGTGACGACGCAGTACGCCCTGGCATCATCCAGGCCC 2033

Qy    2198 AGCCGGACAAGAGCGAGGAGCGGAGCTGGTGAACCAGGATCATCGAGCTGATCAAGAAGG 2257
Db    2034 AGCCCGGACAAGAGCCGAGGAGCGAGCTGGTGAACCAGGATCATCGAGCTGATCAAGAAGG 2093

Qy    2258 AGAAGGTGTACCTGAGCTGGGTGCCCGGCCCACAAGGGCATCGGCGGCAACGAGCAGATCG 2317
Db    2094 AGAAGGTGTACCTGGGCCTGGGTGCCCGGCCCACAAGGGCATCGGCGGCAACGAGCAGATCG 2153

Qy    2318 ACAAGCTGGTGAGCAAGGGCATCGGCACCGTGCTGTTCCTGGACGGCATCGATGGCGGCA 2377
Db    2154 ACAAGCTGGTGAGCGCGCGGCATCGGCCAAGGTGCTGTTCCTGAACGGCCGGCATCGATGGCGGCA 2213

Qy    2378 TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGGCAGCCCGTGAAGATGAATT 2437
Db    2214 TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGGCAGCCCGTGAAGGATGAATT 2273

Qy    2438 AAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2469
Db    2274 AAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2305
```

RESULT 6
US-09-475-515-81
; Sequence 81, Application US/09475515A
; Patent No. 6602705
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: SRIVASTAVA, Indresh
; APPLICANT: LIAN, Ying
; APPLICANT: HARTOG, Karin
; APPLICANT: LIU, Hong
; APPLICANT: GREER, Catherine
; APPLICANT: SELBY, Mark
; APPLICANT: WALKER, Christopher
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
; TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
; FILE REFERENCE: 1621.002
; CURRENT APPLICATION NUMBER: US/09/475,515A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 2299
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: FS(+).proinact.RTopt.YMWM
US-09-475-515-81

Query Match   74.2%; Score 1831.2; DB 4; Length 2299;
Best Local Similarity 88.6%; Pred. No. 3.1e-294;
Matches 2049; Conservative   0; Mismatches 238; Indels 25; Gaps   5;

```
Qy     170 GCGGCAAGGAGGGCCACCGATGAAGGACTGCACCCGAGCGCCACCTCTTCCCGCG 229
Db       1 GCGGCCGCGAAGGACACCAGCCAAATGAAAGATTGCACTGAGAAGAATTGCAGCTAATTTTTT-AGGG 59

Qy     230 AGGACCTGGCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCAGCGAGCAGAAACCGCGCCA 289

```
Qy       1358 GGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGA 1417
                ||      |||||||||||||||||||||||||||||||||||||  |||||||||||
Db       1194 CCAT------CGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCATGCTGCCCGAGA 1247

Qy       1418 AGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCC 1477
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 AGGACAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCC 1307

Qy       1478 AGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCC 1537
              |||||||| ||||||||||||||   |||||||||||||||||||||||| |||||||||
Db       1308 AGATCTACGCCGGCATCAAGGTGAAGCAGCTGTGCAAGCTGCTGCGCGGCACCAAGGCCC 1367

Qy       1538 TGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGA 1597
              |||||||  |  | ||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 TGACCGAGGTGATCCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGA 1427

Qy       1598 TCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGA 1657
              |||||   |||||||||||||  |||||||||||||||||||||||||||||||||||||
Db       1428 TCCTGAAGGAGCCCGTGCACGAGGTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGA 1487

Qy       1658 TCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACC 1717
              |||||||||||||||| | |||||||||||||||||||||||||||||||||||||||||
Db       1488 TCCAGAAGCAGGGCCAGGGCCAGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACC 1547

Qy       1718 TGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACACCAACGACGTGAAGCAGCTGA 1777
              ||||||||||||||||||||   ||||||  |||||||||||||||||||||||||||||
Db       1548 TGAAGACCGGCAAGTACGCCCGCATGCGCGGCGCCCACACCAACGACGTGAAGCAGCTGA 1607

Qy       1778 CCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGT 1837
              ||||||||||||||||| |   ||  ||||||||||||||||||||||||| ||||||||
Db       1608 CCGAGGCCGTGCAGAAGGTGAGCACCGAGAGCATCGTGATCTGGGGCAAGATCCCCAAGT 1667

Qy       1838 TCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCA 1897
              ||   ||||||||||||||||||||||||||| |||||||||  || |||||||||||||
Db       1668 TCAAGCTGCCCATCCAGAAGGAGACCTGGGAGGCCTGGTGGATGGAGTACTGGCAGGCCA 1727

Qy       1898 CCTGGATCCCCGAGTGGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGC 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1728 CCTGGATCCCCGAGTGGAGTTCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGC 1787

Qy       1958 TGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCG 2017
              ||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||||
Db       1788 TGGAGAAGGAGCCCATCGTGGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCG 1847

Qy       2018 AGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGCCGGCAGAAGATCGTGAGCC 2077
              |||||||| | ||||||||||||||||||||||||||||||||||||||| | ||||||
Db       1848 AGACCAAGCTGGGCAAGGCCGGCTACGTGACCGACCGGGGCCGGCAGAAGGTGGTGAGCA 1907

Qy       2078 TGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACA 2137
              |  |||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       1908 TCGCCGACACCACCAACCAGAAGACCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACA 1967

Qy       2138 GCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC 2197
              |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1968 GCGGCCTGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCC 2027

Qy       2198 AGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGG 2257
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db       2028 AGCCCGACAAGAGCGAGAGCGAGCTGGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGG 2087

Qy       2258 AGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCG 2317
              ||||||||||||||  |||||||||||||||||||||||||||||||||||||||| | |
Db       2088 AGAAGGTGTACCTGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGGTGG 2147

Qy       2318 ACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCA 2377
              ||||||||||||||   |||||||||||||||||||||||| ||||||||||||||||||
Db       2148 ACAAGCTGGTGAGCGCCGGCATCCGCAAGGTGCTGTTCCTGAACGGCATCGATGGCGGCA 2207

Qy       2378 TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCGGCCCTAGGATCGATT 2437
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2208 TCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCGGCCCTAGGATCGATT 2267

Qy       2438 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2469
              ||||||||||||||||||||||||||||||||
Db       2268 AAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2299

RESULT 7
US-09-552-950-2
; Sequence 2, Application US/09552950
; Patent No. 6541248
; GENERAL INFORMATION:
;  APPLICANT: Oxford Biomedica (UK) Limited
;  TITLE OF INVENTION: Anti-Viral Vectors
;  FILE REFERENCE: 674524-2004
;  CURRENT APPLICATION NUMBER: US/09/552,950
;  CURRENT FILING DATE:  2000-04-20
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4307
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:gagpol-SYNgp - codon
;   OTHER INFORMATION: optimised gagpol sequence
US-09-552-950-2

  Query Match             68.0%;  Score 1678.6;  DB 4;  Length 4307;
  Best Local Similarity   82.9%;  Pred. No. 5.3e-269;
  Matches 1968;  Conservative    0;  Mismatches  389;  Indels   16;  Gaps    4;

Qy         12 CATGGCCGAGGCCATGAGCCAGG---CCACCAGCGCCAACATCCTGATGCAGCGCAGCAA 68
              | |||| ||||||||||||||||   ||| |  ||| | | |||| |||||||||| ||||
Db       1086 CCTGGCTGAGGCCATGAGCCAGGTGACCAACTCCGCTACCATCATGATGCAGCGCGGCAA 1145

Qy         69 CTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGC 128
              |||   |  ||        ||| |||||||||||||||||| ||||| || || ||||  ||
Db       1146 CTTTCGGAACCAACGCAAGATCGTCAAGTGCTTCAACTGTGGCAAAGAAGGGCACACAGC 1205

Qy        129 CCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCA 188
              |||||||||| | |||||  | || |||||||||||||| ||||||||||| ||||||||
Db       1206 CCGCAACTGCAGGGCCCCTAGGAAAAAGGGCTGCTGGAAATGCGGCAAGGAAGGCCACCA 1265

Qy        189 GATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCA 248
              |||||| ||||| || ||| | ||||| || || ||  | || || ||||||||||  ||
Db       1266 GATGAAAGACTGTACGGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCTACA 1324

Qy        249 GGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGA 308
               || |||||| | || ||  |    ||||||| ||||||||||||||||||| |||||||
Db       1325 AGGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCCGCGCCAACAGCCCCACCCGCCGCGA 1384

Qy        309 GCTGCAGGT------GCGCGGCGACAACCCCCGCAGCGAGGCCGGCGCCGAGCGCCAGGG 362
              ||| |||||       |||| | ||||| ||| |  ||| || || |||| ||||||||
Db       1385 GCTTCAGGTCTGGGGTCGCGACAACAACTCCCCCTCCGAAGCAGGAGCCGACCGCCAGGG 1444

Qy        363 CA------CCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAA 416
              ||      || | ||||||||| ||| |||| || |||||||||| ||||| || | ||||||
Db       1445 CACGGTGTCCTTCAACTTCCCTCAGGTCACGCTTTGGCAGCGACCCCTCGTCACCATCAA 1504

Qy        417 GGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGA 476
              | | || || ||| |||||||||| || || ||||||||| || ||||||||||||||||||
Db       1505 GATCGGGGGGCAGCTCAAGGAGGCTCTCCTGGACACCGGAGCAGACGACACCGTGCTGGA 1564

Qy        477 GGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCAT 536
              |||||||     |||| |||   |||||||| ||||||||||| || || |||||||| |||||
Db       1565 GGAGATGTCGTTGCCAGGCCGCTGGAAGCCGAAGATGATCGGGGGAATCGGCGGTTTCAT 1624

Qy        537 CAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCAC 596
              ||||||||||||||| ||||||||||| ||||| |||||||||| | ||||| ||||| ||
Db       1625 CAAGGTGCGCCAGTATGACCAGATCCTCATCGAAATCTGCGGCCACAAGGCTATCGGTAC 1684

Qy        597 CGTGCTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGG 656
              ||||||| | |||||||| ||||| ||||||||||| |||||| || |||| ||| | ||
Db       1685 CGTGCTGGTGGGCCCCACACCCGTCAACATCATCGGACGCAACCTGTTGACGCAGATCGG 1744
```

Qy 1737 CAAGATGCGCACCGCCACCAACGACGTGAAGCAGCTGAGCCGGTGCGAGAAGAT 1796
Db 2825 CCGGATGAGGGGTGCCCACTAACGACGTGAAGCAGCTGAGCCGGTGCGAGAGAT 2884

Qy 1797 CGCCATGGAGAGCATCGTGATCTGGGCAAGACCCCAAGTTCCGCCTGCCATCCAGAA 1856
Db 2885 CACCACCGGAAAGCATCGTGATCTGGGAAAGACTCCTAAGTTCAAGCTGCCATCCAGAA 2944

Qy 1857 GGAGACCTGGAGACCTGGTGCACCTACTGGTGCGACCACTGGACCTGGATCCCCGAGTGGGA 1916
Db 2945 GGAAACCTGGGAAACCTGGTGGCACAGATATTGGCAGGCCACCTGGATTCCTGAGTGGGA 3004

Qy 1917 GTTCCTGAACACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGAAGAGGAGGCCCATCAT 1976
Db 3005 GTTCGTCAACACCCCTCCCTGGTGAAGCTGTGGTACCAGCTGGAAGAGGAGGCCCATAGT 3064

Qy 1977 CGGGCCGGAGACCTTCTACGTGACGTGGACGGCGCCGCGCCCAACCGCGAGACCAAGGAGACCACCAACCA 2036
Db 3065 GGGGCCGGAAACCTTCTACGTGACGTGGACAGCAGAGATCGTGAGCCTGACCGAGACCACCAACCA 3124

Qy 2037 CGGCTACGTGACCGACCGCGGGCGCCGGAGAAGATCGTGAGCCTGACCGAGTGAACAT 2096
Db 3125 CGGATACGTCACTAACCGGGGCCAGCAGAGAGTTGTCACCCTCACTGACACCACCAACCA 3184

Qy 2097 GAAGACCGAGCTGCAGAGCCAGTAGGCCCTGCCCTGGCATCATCCAGCGGCCAGTGAACAT 2156
Db 3185 GAAGACTGAGCTGCAGGCCCATTTACCTCGCTTGCAGGACTCGGGCCTCGAGGGTGAACAT 3244

Qy 2157 CGTGACCGACAGCAGCCAGTAGCCGCATCATCCAGCCCGACAAGAGCGAGAG 2216
Db 3245 CGTAACAGACTCTCAGTATGCCCTGGGCATCATTCAAGCCCAGCCAGACCAGAGTGAGTC 3304

Qy 2217 CGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTG 2276
Db 3305 CGAGCTGGTGCAATCAGATCATCGAGCAGCAGCTGATCAAGAAGAGGAAAGGTCTATCGGCCTG 3364

Qy 2277 GGTGCCGCGCCACAAGGGCATCGGCGGCAACGACGAGATCGACAAGCTGGTGAGCAAGGGG 2336
Db 3365 GGTACCCGCCCCAAAGGCATTGGCGGCAATGAGCACGTCGACAAGCTGGTCTCGGCTGG 3424

Qy 2337 CATCCCGCAAGGTGCGTTCCTGGACGGCATCGA 2369
Db 3425 CATCAGGAAGGTGTATTCCTGGATGGCATCGA 3457

Query Match 66.9%; Score 1651.4; DB 4; Length 9772;
Best Local Similarity 82.2%; Pred. No. 1.8e-264;
Matches 1951; Conservative 0; Mismatches 406; Indels 16; Gaps 4;

Qy 12 CATGGCGAGCCATGAGCCAGG---CCACCAGCCGCCAACATCCTGATGCAGCCAGCAA 68
Db 2193 CCTGGCTGTGAGGCCATGAGCCAGGTGACCAGCCAGGTGACCAACTCCGCTACCATCATGATGCAGCCAGCCAGCGCGGCAA 2252

Qy 657 CTTGCACCCTGGAACTTCCCATCAGCCCATCGAGACCTGCCGTGAAGCTGAAGCCCGG 716
Db 1745 TTTGCACGTGACCTTCCCCATTAGCCCTATCGAGACCGGTACCCGTGAAGCTGAAGCCCGG 1804

Qy 717 CATGGACGGCCCCAAGGTGAACGCAGTGCCCCTGACCGAGGAGGAAGATCAAGGCCCTGAC 776
Db 1805 GATGGGACGCCCGAAGGTCAACATTGGCCATTGACAGAGGAGGAAGATCAAGGCCACTGGT 1864

Qy 777 CGCCATCTCGAGGAGATGGAGAAGGAGGGCAAGATCAGCCCCCGAGACCC 836
Db 1865 GGAGATTTGCACAGAGATGGAAAAGGAAGGGCAAAAATCTCCAAGATTGGGCCTGAGAACCC 1924

Qy 837 CTACAACACCCCCGTGTTGTTCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGTGGT 896
Db 1925 GTACAACACGCCCGTGTTCGCAATCAAGAAGAAGGACTCGACGAAATGGCGCAAGCTGGT 1984

Qy 897 GGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCCCTA 956
Db 1985 GGACTTCCGCGAGCTGAACAAGCGCACGCAAGACTTCTGGGAGGTTCAGGTGGGCATCCC 2044

Qy 957 CCACCCCGCCGGCCTGAAGAAGAAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTA 1016
Db 2045 GCACCCCCGCCAGGGCTGAAGAAGAAGAAATCCGTGACCGTACTGCCTTCACAATCCCTCGAT 2104

Qy 1017 CTTCAGGTGCCCCTGGACGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCCAGCAT 1076
Db 2105 CTTCTCCGTTCCCCTGGACGAAGACTTCAGAGAGTACACTGCCTTCCACAATCCCTTCGAT 2164

Qy 1077 CAACAACAGACCCCCCGCATCCGTACCAGTACAACGTGTCCAGGGCTGTGGAAGGG 1136
Db 2165 CAACAACGAGACCACCCGGGGGATTCGATATCAGTACAACGTGTCCAGGGCTGGAAAAGG 2224

Qy 1137 CAGCCCCAGCATCTTCAGAGCAGCATGACCAAGATCCTGAGCCCTTCCGCGCCCCCGCAA 1196
Db 2225 CTCTCCCCCAATCTTCAGAGTAGCCATGAGTTCAGAGCCAAAATCCTGGAGCCTTTCCGCGCACCTGAGACAGAT 2284

Qy 1197 CCCCGAGATCGTGATCTACCAGTACACCATGACGGTGGGCCAGCGACCTGGGAGATCCTGGAGAGAT 1256
Db 2285 CCCCGACATCGTCATCTATCAGTACATCATGGATACGGTGGGCTCTGATCTAGAGAT 2344

Qy 1257 CGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGGCAAGCACCTGCTCGCGCTGGGGCTTCAC 1316
Db 2345 AGGGCAGCACCGCGCCAAGATCGAGGAGCTGCGGCCAGCACCGTGTTGAGGTGGGACTGAC 2404

Qy 1317 CACCCCGACACAGAAGCACCAGAGAGGAGCGCCCCCCTTCCTGTCGATGGGATCGGCTACGAGCTGCA 1376
Db 2405 CACACCCCGACACAGAAGACCAGAGAGGAGCCCTCCCTTCCCTCCTTCCTGATGGATGGGTTACGAGCTGCA 2464

Qy 1377 CCCCGACAAGTGGACCGTGCAGCCATCGAGCTGCCCGAGAGGAGAGCTGGACCGTGAA 1436
Db 2465 CCCTGACAAATGACCGTGCAGCCTATCGTGTCTGCCAGAGAGAAGACAGCTGGACTGTCAA 2524

Qy 1437 CGACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCCAGCAGATCTACCCCGGCATCAA 1496
Db 2525 CGACATACAGAGCTGGTGGTGGGAAGTTGAACTGGGCCAGTCAGATTTACCCAGGATTAA 2584

Qy 1497 GGTGCGCCAGCTGTGCAAGCTGCTGCGCGGGCCGGCCAAGGCCCTGACCGACATCGTGCCCT 1556
Db 2585 GGTGAGGGAGCTGTGCAAACTCCGCGGGGCAAGCAAGGCACTCACAGAGGTGATCCCCT 2644

Qy 1557 GACCGAGGAGGCCGAGCTGGAGCTGGCAGAAAACCGCGAGAATCCTGCGCGCAGCCGTGCA 1616
Db 2645 AACCGAGGAGGCCGAGCTCGAACTGGCAGAAAACCGGAGATCCCAGAAGGAGCCCCGTGCA 2704

Qy 1617 CGGCGTGTACTACGACCCCGAGAAGGAGACCTGTTGGTGGGCCAGAATCCAGAAGACAGGGCCACGA 1676
Db 2705 CGGCGTGTACTATGACCCCTCCAAGGACCTGATCGCCGAGAATCCGGGAGAAGCACAGCTGCAAGG 2764

Qy 1677 CCAGTGGACCTACGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGC 1736
Db 2765 CCAGTGGACCTATTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGC 2824

```
Qy    69  CTTCAAGGGCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGC  128
          |||||         ||||||| ||||| ||| ||   ||||||||   ||| |||||||
Db  2253  CTTTCGGAACCAACGCAAGATCGTCAACTGCGTTCAACTGTGGCAAGAAGGGCACACAGC  2312

Qy   129  CCGCAACTGCGCCGCCCGCCGCCCTAGGAAAAAGGCTGTTGAAAATGTGGAAGGACACCA  188
          ||||        |||  || ||||||  ||    ||| |  ||||||  |||||  ||||
Db  2313  CCGCAACTGCAGGGCCCCTAGGACAGAGGGCTGTTGAAAATGTGGAAAGGAAGGACACCA  2372

Qy   189  GATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCA  248
          |||||||      |||||||||||||| ||||||||||  ||||||| |||||||||||
Db  2373  AATGAAAGATTGTACTGAGGAGACAGGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCCACA  2431

Qy   249  GGGCAAGGCCCCGCCAGTTCCCAGCAGCAGAACCGGCCAACAGCCCACCAGCCGCGA  308
          |||        | ||||||  |   |  |   | ||||  ||||  | |  ||||||
Db  2432  AGGGAAGGCCAGGGAATTTTCTTCAGCAGACCAGGCGACCAGCCAACAGCCCACCAGAGAGA  2491

Qy   309  GCTGCAGGTGCGCGG------CGACAACCCCGGAGCGAGGCCGGCGCCGAGCGCCAGGG  362
          | | | ||  | ||       ||  |  |||  ||| | |  ||| || |  || |||
Db  2492  GCTTCAGGTTTGGGAGAGACAACAACTCCCTCTCAGAAGCAGGAGGCGATAGACAAGG  2551

Qy   363  CA------CCCTGAACTTCCCCAGATCACCCTGGCAGCCCAGGACGCCCCTCGTCACAATCAA  416
          ||       | |  | |||||||||| |   ||    | | |    | |||| |||||||||
Db  2552  AACTGTATCCTTTAGCTTCCTCAGATCACTCTTTGGCAGCCGGACCACACCGTCTGTCACAATAAA  2611

Qy   417  GGTGGCGGCCCAGATCAAGAACGAGAGGCCCTGCTGGAAGCCCAAGATGATCGGCGCGCTTCAT  476
          ||||   ||||| |||||  || || || ||   |||||   | | |  ||||| |||||||||
Db  2612  GATAGGGGGGCAGCTCAAGGAGGCTCTCCTGGAAGCCAACACCAACCGTGCTGGAATCGGCGGTTTCAT  2671

Qy   477  GGAGATGAGCCTGCCCGGCAAGTGAAGCCAAGATCTGCGCGGAAGAGGGCCATCGGCGCAC  536
          ||||| |||| |||||||||   | | ||   |  |||| | ||| || ||| | | ||
Db  2672  GGAGAGTGCGTTGCCAGGCCGCTGGAAGCCGAAGATTGACAGAGGCCGTATCGGCGTATCGGTAC  2731

Qy   537  CAAGGTGCCCAGTACGACACCAGAGATCTGCGCCAACATGCTGACCCAGCTGGG  596
          ||||||||||||| |||| |  |  |||| ||   |||| |||| || |  |||
Db  2732  CAAGGTGCCCAGTATGACCAGATACTCCGGCGCCAACAGGCTATCGGTGGG  2791

Qy   597  CGTGCTGATCGGCCCCACCCGTGAACATCATCGGCCGCCAACATGCTGACCCAGCTGGG  656
          |||||| |  |||  || | |   |||| |||| || || ||| | ||| ||||||||
Db  2792  CGTGCTGGGTTGGGGCCCCACCACCCGTCAACATCGGACGCAGATCGG  2851

Qy   657  CTGCACCCTGAACTTCCCCATCAGCCCCATCGAGACCTGCCCGTGAAGCTGAAGCCCGG  716
          || ||||||||| |||||||| ||  ||| |||| || |  | || |||||||||||||
Db  2852  TTGCACGTCGAACTTCCCATTAGCCGTCTATCGAGAGACGTCGGTGAAGCTGAAGCCCGG  2911

Qy   717  CATGGACACGGCCCAAGGTGAAGCAGTGGACCTTCTGGACGAGGAGATCAAGGCCCTGAC  776
          ||| ||||| ||||||||||||| ||   |  | || || |||| |||| ||| |  ||
Db  2912  GATGGACACGGCCCGAAGGTCAAGGAGGCCATTGACAGAGGAGAAGATCAAGGCCACTGGT  2971

Qy   777  CGCCATCTGCAGAGGAGGATGGAGAAGAGGGCCAAGATCACCAAGATCGGCCCCGAGACCC  836
          ||   | | ||  | ||||  ||||| | |  |||| |  |  |||| |||  ||||| ||
Db  2972  GGAGATTTGCACAGATCGGCGAAAGGAGAACGGAAAATCTTCCAAGATTGGGCCTGAGAACCC  3031

Qy   837  CTACAACACCCCGTGTTCGCCATCAAGAAGAGGAGCAGCACCAAGTTGGCGCAAGCTGGT  896
          || |||||||||||| | || || ||  || |||| || || ||   ||||   |||||
Db  3032  GTACAACAGCCGTGTTCGCCAATCAAGAGGACTCGACGAAATGCGCAAGGT  3091

Qy   897  GGACTTCGGCGCCTGGAACAGAGGCCACCCCAGAGGACTTCTGGAGGTGCAGCTGCGAGCTCCC  956
          ||| || ||||| ||||||| ||||| || || |||| |||| || || |  ||| | ||||
Db  3092  GGACTTCGGCGCCTGAACTTCACGGTTCTGGGAGGTTCAGCTGGATGTGGGCCATCCC  3151

Qy   957  CCACCCCGCCGCCCAGGGCACGCCGACTACTGGATGTGGGGTGTCATGGCCGCCTA  1016
          ||   |||| |||||||  || |  |||| |||| | || ||||||||  |||||||
Db  3152  GCACCCGCAGCATCTTCAGAGCAGCATCCGTGACCAAGCAAATGCGAGCCTA  3211

Qy  1017  CTTCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCCGCCTTCACCATCCCCAGCAT  1076
          ||| |   |||||  |||||| ||||||||||  || ||||||| || |  ||||  ||
Db  3212  CTTCTCCGTTCCCCTGGACGAGGACTTCAGGAGCGTACACCGGCCTTGCGCCTTCACAATCCCTCGAT  3271

Qy  1077  CAACAACGAGACCCCCGGCATCTCCGCTACCAACGGTCTGCCCCCAGGGCTGGAAGGG  1136
          |||||||||| || ||||| |||| | ||| ||||   |  || | |||||||| |||
Db  3272  CAACAACGAGACACCGGGGATTCGATATCAGTCAATATCGTTTCAACGTGCTTCCCCCAGGGCTGGAAGG  3331

Qy  1137  CAGCCCCAGGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCGCCCGCAA  1196
```

```
Db  3332  CTCTCCCGCCAATCTTCCAGAGTAGCAGTGACGATGACCAAAAATCCTGGAGCCTTCCGCAAACAGAA  3391

Qy  1197  CCCCGAGATCGTGATCTACCAGTGACAGTACATGGACGACCTGTAGTGGGGCAGCGACCTGGAGAT  1256
Db  3392  CCCGACATCGTCATCATCATCAGTACATGAGATGACTTGTACGTGGGTCTGATCTAGAGAT  3451

Qy  1257  CGGCCAGCAGCACCACCGCCAAGAGCTGCGCAAGAGCTGCGCAAGCACCTGCTGCGGCTTCAC  1316
Db  3452  AGGGCAGCACCGGCACCCAAGACGAGCGCCAGCACCTGTTGAGGTGGGACTGAC  3511

Qy  1317  CACCCCCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGTGGGATGGCTACGAGCTGCA  1376
Db  3512  CACACCCGACAAGAAGCACCAGAAGGAGCCCCCTTCCTGGGATGGGTACGAGCTGCA  3571

Qy  1377  CCCCGACAAGTGACCGTGCAGCCATCGAGCTGCCCGAGTGCCCGCAGAAGGAGAGCTGGACCGTGAA  1436
Db  3572  CCCTGACAAATGACCCTGCAGCCTATCGTGCTGCGCAGGAACAGCTTGGGACTGTCAA  3631

Qy  1437  CGACATCCAGAGCTGTGCAAGTCTGAACTGGCGCGGCGAGATCTACCCGGCATCAA  1496
Db  3632  CGACATACAGAGCTGTGTGCAAACTCTCCGGGAACTGGGCCAGTTGAACTGGCCAGTCAGATTTACCCAGGGATTAA  3691

Qy  1497  GGTGCGCCAGCTGTGCAAGTGCTGTGCGGCCCCGAGATCCTGGCCCCTGACCGACATCGTGTCCCCCT  1556
Db  3692  GGTGAGGCAGCTGTGTGCAAACTCTCCGGAACTGGGCACTCACAGAGGTGATCCCCT  3751

Qy  1557  GACCGAGGAGGCCGAGCTGGAGAGCTGGCCGAGAGATCTTGCGCGGAGCCCGTGCA  1616
Db  3752  AACCGAGGAGGCCGAGCTGCGAACTGGAGAGAAAACCGAGAGATCCTAAAGGAGCCCGTGCA  3811

Qy  1617  CGGCCGTGACTACGACCCCAGCAAGAACCTGGTGGCGAGATCCAGAAGCAGGGCCACGA  1676
Db  3812  CGGCGTGACTATGACCCCTCCAAGATCGCCGAGATCCAGAAGCAGGGGCAAGG  3871

Qy  1677  CCAGTGACCTACCAGATCTACCAGGAATCTACCAGGACCCTTCAAGAACCTGAAGACTACGC  1736
Db  3872  CCAGTGACCTATCAGATTTACAGGGACCCTTCAAGAACCTGAAGACCGCAAGTACGC  3931

Qy  1737  CAAGATCGCACCGCCACCACCAACCAGTGAAGACCCCCCCAAGTTCCGCCTGCCCATCCCAGAA  1796
Db  3932  CCCGCAATGAGGGCGTCCCCACCACTAACGACGTCAAGCAGTGTGTGACCCAGAAGAT  3991

Qy  1797  CGCCATCGAGAGCATCGATCTGGGGACCTGATCTGGACCTACTACTGGACCGAGTGGGA  1856
Db  3992  CACCACCGAAAGGCATCGAGTCGTGATCGTGGGGAAAGACTCCTAAGTCAAGCTGCCAGAA  4051

Qy  1857  GGAGACCTGGGACCTGGTGGACCCACTGGATCCCCCGAGTGGGA  1916
Db  4052  GGAAACCTGGGAAAACCTGGTGACAGAGTATTGGCAGGCCACCTGGAGTGGGGA  4111

Qy  1917  GTTCGTGAACACCCCCCCCTCCCCTGGTGAAGCTGTGGTTACCAGTGAAGCTGGAGAAGGAGCCCATCAT  1976
Db  4112  GTTCGTGTCAACACCCCCTTCCCCTGGTGAAGCTGTGGTTACCAGTGAAGGAGCCCATAGT  4171

Qy  1977  CGGCTACGTGACGACCCTGCAGGCCATCAGCTGCCCTGAGGACGCAGGGCCGAGACCCAAGGC  2036
Db  4172  GGGCGGCGAAACTTCACGGTGGATGGGCCGCTGAGGGCCAACTAAGCTGGGCAAAGC  4231

Qy  2037  CGGCTACGTGACCGACGTGCCCTGAGGATCGTGAGGCTGACCGAGAACACCAACCA  2096
Db  4232  CGGATACGTCACTAACGGGGCAGCAGAAGGTTGTCACCCTCACTGACACACCAACCA  4291

Qy  2097  GAAGACCGAGCTGCAGACCTCAGCTGCCCTGCAGGACAGCGGCGAGTGAACAT  2156
Db  4292  GAAGAGGACTGAGCTGCAGGGCCATTTACCTCGCTTCGCTTGCAGGAGGTGAACAT  4351

Qy  2157  CGTGACCGAGCAGCCAGTACGCCGCATCATCCAGGGCAAGAGAGCCGAGAG  2216
Db  4352  CGTGACAGACTCTCAGTTCATGTCCCGCATTCAACGCCCCAGCAGCCAGTGAGTC  4411

Qy  2217  CAGCTGGTGAACCAGATCATCGAGCGCTGATCAAGAGAGGTGTTACCTGAGCTG  2276
```

```
Db   4412  CGAGCTGGTCAATCAGATCATCGAGCAGCTCGATCAAGAAGGAAAAAGGTCTATCTCGGCCTG 4471
Db   2276  TCGGGGGCAACTCAAGGAAGAGGCTGCTCGATCAGGAGCAGATGATCAGATTAGAAG 2335

Qy   2277  GGTGCCCGCCCACAAGGCATCGGCGCGGCAACGAGACGACAAGCTGGTGAGCAAGGG 2336
Qy    479  AGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGACCCACGGCCATCGGCGCTTCATCA 538

Db   4472  GGTACCCGCCCCACAAAGGCATTGGCGCGGCCAATGAGCAGGTCCACAAGCCTGGTCTCCGCTGG 4531
Db   2336  AAATGAGTTTGCCAGGAGATGGAAACCAAAATGATAGGGGGATCGGGGGCTTCATCA 2395

Qy   2337  CATCCGCGACAAGGTGCTGTTCCTGGACGGCATCGA 2369
Qy    539  AGGTGCGCGACCAGTACGACCAGATCCTGATCGAGATCTGCGGCCAAGAGGCCATCGGCCACCG 598

Db   4532  CATCAGGAAGGTGCTATTCCTGGATGGCATCGA 4564
Db   2396  AGGTGAGGCAGTACGACCAGATACTCATAGAAATCTGTGGACATAAGCTATAGGTTACAG 2455


RESULT 9
US-09-872-733A-6
; Sequence 6, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
;  APPLICANT: The Government of the United States of America, as
;  TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
;  TITLE OF INVENTION: SIV ENV GENES
;  FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
;  CURRENT APPLICATION NUMBER: US/09/872,733A
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: PCT/US00/34985
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/173,036
;  PRIOR FILING DATE: 1999-12-23
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 6
;  LENGTH: 8366
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
;  OTHER INFORMATION: of the construct pCMVgagpolBNKan containing a CMV
;  OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
;  OTHER INFORMATION: resistance gene
US-09-872-733A-6

Query Match            64.5%;  Score 1592.8;  DB 4;  Length 8366;
Best Local Similarity  80.7%;  Pred. No. 8.3e-255;
Matches 1914;  Conservative   0;  Mismatches  442;  Indels  16;  Gaps   4;


Qy    14  TGGCCGAGGCCATGAGCCAGGCCACCAGC--GCCAACATCCTGATGCAGCGCAGCAACT 70
Db  1657  TGGCCGAGGCGATGAGCCAGGTGACGAACTCGGCGACTCGGCCGATAATGATGCAGAGAGGCAACT 1916

Qy    71  TCAAGGGCCCCAAGGCCGATCATCAAGTGCTTCAAGTGCTTCAACTGCGGCAAGAGGGCCACCATCGCCC 130
Db  1917  TCCGGAACCAGCAGCGGCGGAAGATCGTCAAGTGCTTCAATTGTGGCAAAGAAGGGCACACCGCCA 1976

Qy   131  GCAACTGCCGCGCAGAGCGCCAGGCCAACTTCTTCCGCGGAAGTCGCAAGAAGTCGGCAAGGGCCACCAGA 190
Db  1977  GGAACTGCCGGGCCCCCGGCCCCGGAAGAAGGCTGTTCAGAAATGTGGAAATGGGAAGGAAGGACACCCAA 2036

Qy   191  TGAAGGACTGCACGAGCGCAGGCCAACTTCTTCCGCGGAAGTCGCAGAGAGACAGGCTAA-TTTTTTAGGGAAGCTC 250
Db  2037  TGAAAGATTGTACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCTTCCTACAAG 2095

Qy   251  GCAAGGGCCCGCGAGTTCCCCAGCGAAGACAGAACCGGCGCCAACAGCCCCACCAGCCGCGGCGAGC 310
Db  2096  GGAAGGCGCAGGGAATTTTCTTCAGCAGCAGCAGCCAACAGCCAACAGCCCCACCAGCCGAAGAGAGC 2155

Qy   311  TGCGGTGCGCGG------CGACAACCCCGCAGCGAGGCCGGCCCGCCGAGGCCAGGGCA 364
Db  2156  TTCAGGTCTGGGTAGAGACAACAACTCCCCCTCAGAAGCAGGAGCCGATAGACAAGGAA 2215

Qy   365  -----CCCTGAACTTCCCCAGATCACCCTGGCAGCGCCCCGTCGTGGTGAGCATCAAGG 418
Db  2216  CTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCGTCACGGTGCCCCTGA 2275

Qy   419  TGGCGGCCAGATCAAGGAGGCCCGTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGG 478
Db  3356  TTAGGCAGCTGTGCAAGCTGTTCGAGGAACCAAGGCACTGACAGAAGTGATCCCACTGA 3415
```

; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733A-1

Query Match   63.1%;   Score 1557.4;   DB 4;   Length 4338;
Best Local Similarity   80.5%;   Pred. No. 5.4e-249;
Matches 1914;   Conservative   0;   Mismatches 441;   Indels   22;   Gaps   7;

```
Qy       14 TGGCCGAGGCCATGAGCCAGGCCACCAGC---GCCAACATCCTGATGCAGCGCAGCAACT 70
            ||||||||| ||||||||||||||||||||   ||||||||||||||||||||||||||||
Db     1085 TGGCCGAGGCGATGAGCCAGGTGACGAACTCGGCGACCATAATGATGCAGAGAGCCAACT 1144

Qy       71 TCAAGGGCCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCC 130
            ||||||||||||||||||||| |||||||||| |||||||||||||| ||||||||||||
Db     1145 TCCGGAACCAGCAGCGGAAGATCGTCAAGTGCTTCAATTGTGGCAAAGAAGGGCACCGCCA 1204

Qy      131 GCAACTGCCGCGCGCCCCCCGCAAGGAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCCACCAGA 190
            ||||||| ||| |||||||| ||||||||||||||||||||||||||||||||||||||||
Db     1205 GGAACTGCCGGGCGCCCCCGGAAGAAGGGCTGTTGGAAATGGGAAAATGGGAAGGAACCAAA 1264

Qy      191 TCGTGAAACCCCCCCCGTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCG 250
            ||||||| |||||||||   ||||||||||||||||||||| |||||||||||||||
Db     1265 TGAAGGACTGACCGAGGCGCCAGTTCTTCCGGAGGAAGGAACCCCATCCTGG--TTTTTAGGGAAGATCTGGCCTTCCTTACAAG 1323

Qy      251 GCAAGGCCCGCGAGTTCCCCAGCGGCGAGAAGAACCGGCCACCAGGCCGAGACCCGCGAGC 310
            ||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db     1324 GGAAGGCCAGGGAATTTTCTTCAGAGCAGCAGACCAGACCAGCCCCAGAAGAGAGC 1383

Qy      311 TGCAGGTGCGCGG------CCACAACTTCCCCGATCCTCTTGGCAACGGCCAGGGCA 364
            ||||||||||||||      ||||| |||||||||||| ||||||||||||||||||
Db     1384 TTCAGTCTCTGGGTAGAGACAATCAGCCCTGACCCTGACTCACTCCGTCGTCGTGAGCATCAAGG 1443

Qy      365 ------CCTGAACTTCCCCCTCTTTAACTTCCCTCGATCCTCAGATCCTCTTGGCAACGGCCAGGGA 418
                  |||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     1444 CTGTATCCTTTAACTTCCCTCAAGGAGGCCCTGCCCCGAGCCCCGAGAGCGAGAGCG 1503

Qy      419 TGGGCCGGCCAGATCAAGGACCCCCTGCCGGACACCGGCCGAGACCAATCCGAGTCAG 478
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1504 TCGGGGGCCAACTCAAGGAGGCGGCTGCTGCTCGATACAGCGACGCAGATGATACAGTATTAGAAG 1563

Qy      479 AGATGAGCTGCCCAAGTTGCGAGCCCAAGATGTGGGCGCATCGGCGCGCTTCATCA 538
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db     1564 AAATGAGTTTGCCAGAGGATCCGCGGCCAACCAAGAAATGATAGGGGGGCTTCATCA 1623

Qy      539 AGGTGCGCCAGTACGACCAGAACCTGATCGACATCGAGCATTTGCGCGGAGGCATCGGCCACCG 598
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     1624 AGGTGAGGCAGTACGACCAGACCAGACCAGATCATAGAAATCGTGTGACATAAGCTATAGGTACAG 1683

Qy      599 TGCTGATCGGCCCCACC----CCGTGAACATCATCGCCCAACATGCTGACCCAGCTG 654
            |||||||||||||||    ||||||||||||||| |||||||||||||||||||
Db     1684 TATTAGTAGGACCTACCTGCAACACTGTCAACATAATTGAAGAAATCTGTTGACCCAGATC 1743

Qy      655 GGCTGCACCTGAACTTCCCCATCAGCCCCATCGAGACCTGCGAGCCC 714
            ||||||||| |||||||||||||||||||||||||||||||||||||
Db     1744 GGCTGCACCTTGAACTTCCCCATCAGCCCTATTGAGACGGTCCCGTGAAGTTGAAGCCG 1803

Qy      715 GGCATGAGCGGCCCCAAGTGAAGCAGTGGCCCGACGGAGGAAGATCAAGGCCCTG 774
            ||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db     1804 GGGATGGACGGCCCCAAGGTCAAGCAATGGCCATTGACGACGCAAGAGATCAAGGCCTTA 1863

Qy      775 ACCGCCATCTGCGACGAGATGGAGAAGGAGGGCAAGATCACCAGGCCAGAACC 834
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1864 GTCGAAATCGTACAGAGATGGAGAAGGAAGATCAGCAAGATCGGGGCTGAGAAC 1923

Qy      835 CCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACACCAGCACCAAGTGCGCCAAGCTG 894
            ||||||||||||||  ||| ||||| |||||||| |||||||||||
```

```
Qy     1559 CCGAGGAGGCCTGGAGCTGGCCAGGAACCGCGAGATCCTGCCGCGAGCCGTGCACG 1618
            ||| |||||| || ||||| |||||||||||||||||||||||||
Db     3416 CAGAGGAAGCAGCAGCTAGAACTGGCAGAGAACCGCGAGAGATCCGAGAGATCCAGGTACATG 3475

Qy     1619 GGCGTGTACTACGACCCCCAGCAAGGACCTGGTGCCGGAGATCCAGAAGCAGAGCCCACGACC 1678
            |||||||||||||||||||||||| |||| ||  ||||||||||||  || ||||||
Db     3476 GAGTGTACTACGACCCCACCCCAGCCAAGCAAGCAGGACCTGATCCGCAGAGATCCAGAGAGGCC 3535

Qy     1679 AGTGGACCTACCAGATCTACCAGAGACCTGAAGACCTGAAGACCTACGCCA 1738
            |||||||||| |||||| |||||||||||||||||||||||||||||||
Db     3536 AATGGACCTACCAAAATCTACCAGAGACCTGAAGACCTTCAAGACCTACGCAA 3595

Qy     1739 AGATGCGCACCGCCCACCAACCACGACGTGAACGAGCAGCTGAAGCAGATGCAGAGATCA 1798
            |||||||||||||||||||||||| ||||| ||||| |||||||||||||||||||
Db     3596 GGATGAGGGGTGCCCACCAACCACCAGATGTGAAGCAGCTGACAGAGGCAGTGCAGAAGATCA 3655

Qy     1799 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCCCCTGCCCATCCAGAAGG 1858
            |||||||||| ||||||||||||||||| |||||||||||||||||||  ||||||
Db     3656 CCACAGAGAGCATCGTGATCTGGGGCAAGACTTCCCAAGTTCAAGCTGCCCATACAGAAGAG 3715

Qy     1859 AGACCTGGGAGACCTGGTGGACCTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1918
            |||||||||||||| |||| ||||||||||||||||||| ||||| |||||||||
Db     3716 AGACCTGGGAGACTGGTGGACCGAGTACTGGCAAGCCACCTGGATCCCTGAGTGGGAGT 3775

Qy     1919 TCGTGAAACCCCCCCCGTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCG 1978
            |||||||  |||||||||   ||||||||||||||| ||||||| ||||||||||
Db     3776 TCGTGAAACACCCCCTCCCTTGGTGAAACTGTGGTATCAGCTGGAGAAGGAACCCATCCTGG 3835

Qy     1979 GCGCCGAGACCTTCTACGTGGACGGCGCGCCCAACCCGAGACCCAACATCGGCAAGGCCG 2038
            ||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||
Db     3836 GAGCAGAGACCTTCTACGTGGATGGGGCAGCCAACAGGGAGACCAAGCTGGGCAAGGCAG 3895

Qy     2039 GCTACGTGACCGACCGGGGGCCGGCCAGAGATCGTGACCTGAGCCTGAGCCTGACCGAGACCAACCAGA 2098
            ||||||||||||||||||  ||||| ||||||| |||| || |||| ||||||||||||
Db     3896 GCTACGTGACCAACCGAGGACAGACGAAAGTGTGACCCTGACTGAC-CACCAACCACCAGA 3955

Qy     2099 AGACCGAGCTGCAGGCCATCTACCTGGCGCCCCTGAGCGTGAAGCGCGGCGAGGTGAACATCG 2158
            ||||||||||| ||||||||||| || || || ||||| ||||||||||||||||||||||
Db     3956 AGACTGAGCTGCAAGCCATCATCCTGGCCCTGCAAGACCAGCCGAGTGAAGTGAACATCG 4015

Qy     2159 CTGTATCCTTTAACTTCCCTCAAGGCCCTGAGCACCCGAGCAAGAGAGCGAGAGCG 2218
            ||||||  || ||||||||| |||||||| || ||||||||||||||||||||||
Db     4016 TGGGCCGGCCAGATCAAGGAGGCCCTGCCGGACACCGGCCAATCCGAGTCAG 4075

Qy     2219 AGTGGTGAACCAGATCATCGAGCTGATCAAGAAGGAAGATGTGACCTGAGCCTGGG 2278
            || |||||||| ||||| ||||||||| |||||||||||| || ||||||||
Db     4076 AGTTGGTGAACCAGATCATCGAGCTGATCAAGAAGGAGAAAGTGTCACCTGGCATGGG 4135

Qy     2279 TGCCGCCCACAAGGGCGATCGGCGGCCAACGAGCAGATCGACAAGCTGCGTGAGCAAGGGCA 2338
            |||||||||||||| || ||||||||||||||||||||||||||||||||||||
Db     4136 TACCAGCACCAAAGGAATTGGAGGAAATGGGGCAGAAACAAGAGTAGTAAATTAGTCAGTGTCTGGGA 4195

Qy     2339 TCCCGCAAGGTGCTGTTCCTGGACCGCATCGAT 2370
            ||||||||||| ||||| |||||||| ||||||
Db     4196 TCCCGGAAGGTGCTGTTCCTGGACCGGGATCGAT 4227
```

RESULT 10
US-09-872-733A-1
; Sequence 1, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 1999-12-23

Db   1924 CCCTACAACACTCCAGTCTTGCAATCAAGAGAAGGACAGTACCAAGTGGAGAAAGCTG 1983
Qy   895 GTGGGACTTCCGCGAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGCATC 954
Db   1984 GTGGGACTTCAGAGAGTGAACAAGAGAACTCAGGACTTCTGGGAAGTTCAGTCGGGCATC 2043
Qy   955 CCCCACCCCGCCGGCGTGAAGAGAAGCGTGAAGAGCGTGCTGGACGTGGGGCGACGCC 1014
Db   2044 CCACATCCGCGTGGGTTGAAGAGAGAGAAGAAGTCAGTGACAGTGCTGGATGTGGGTGATGCC 2103
Qy   1015 TACTTCAGCGTGCCCCTGGACGAGAGACTTCCGGACAGACTACACCGGCTTCACCATCCCCAGC 1074
Db   2104 TACTTCTCCGTTCCCTTCGGACGAGGACTTCAGGAGTACACTGCCTTCACGATACCTAGC 2163
Qy   1075 ATCAACAACAAGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAG 1134
Db   2164 ATCAACAACAAGAGACACCCAGGCATCCGCTACCAGTACAACGTGCTGCCCCACAGGGATGGAAG 2223
Qy   1135 GGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGC 1194
Db   2224 GGATCACCAGCCATCTTCAAAGCAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAA 2283
Qy   1195 AACCCCGAGATGTGATCTACCAGTACCATGACATGGACGACCTGTACGTGGCAGGGACCTGGAG 1254
Db   2284 AACCCAGACATCGTGATCTATCAGTACATGGACGACCTCTACGTAGGAAGTGACCTGGAG 2343
Qy   1255 ATC-GGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCCGCTGGGGGCTT 1313
Db   2344 ATCGGGGCAGCAGGACCAAGATCGAGGAGCTGAGACGCCTGTCGTTGAGGTGGGGGACT 2403
Qy   1314 CACCACCCCGACAAGAAGCACCAGAAGAAGGAGCCCCTTCCTGTGTGGTGGGCTACGAGCT 1373
Db   2404 GACCACACCAGACAAGAAGCAACCAGAAGAGGAACCTCCTGTGTGGTGGGCTACGAACT 2463
Qy   1374 GCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCCGAGAAGGAGAGCTGGACCGT 1433
Db   2464 GCACCTGACAAGTGGACAGTGCAGCTCGAGCTGCCATCGTGCTGCGAGAAGGACACAGCTGACTGT 2523
Qy   1434 GAACGCAGCATCCAGAAGCTGGTGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCAT 1493
Db   2524 GAACGCAGCATACAGAAGCTCGTGGGCAAGCTGAACTGGGCAAGCCAGATCTACCCCAGGCAT 2583
Qy   1494 CAAGGTGCGCCAGCTGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCC 1553
Db   2584 CAAAGTTAGGCAGCTGCTGTGTCAAGCTGCTTCGACGAGCAACCAAGGCACTGACAGAAGTGATCCC 2643
Qy   1554 CCTGACCGAGGAGGGCGAGCTGGAGCTGGCCGAGAAGAACCGGCGCGAGATCCTGCGCGCCCGT 1613
Db   2644 ACTGACAGAGGAAGCAGGAGCTAGAACTGGCAGAGAGAACCAGAGATCCTGAGAGGAGCCAGT 2703
Qy   1614 GCACGGGCTGTGTACTACGACCCCAGCCAGAGGACCTGGTGCCCAGAGAGCCAGAGGCCA 1673
Db   2704 ACATGGAGTGTACTACGACCACCCAAGCCAAGGACCCTGATCGCGACCAAGGATCCCAGAAGATCCAGAGAGCCAGGGCA 2763
Qy   1674 CGACCAGTGACCTACCAGATCTACCAGGAGACCTTCAAGAACCTGAAGACCGGCAAGTA 1733
Db   2764 AGGCCAATGGACCTACCACCAAATCTACCAGGAGCCCTTCAAGAACCTGAAGACAGAGCACAGGCAAGTA 2823
Qy   1734 CGCCAAGATGCGCACCGCCCACACCCACAACGACGTGAAGCAGCTGAGCGAGGCCGTGCAGAA 1793
Db   2824 CGCAAGGATGAGGGGGTGCCCCACCACCACAACGATGTGAAGCAGCAGTGCAGAA 2883
Qy   1794 GATCGCCGCCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCATCCA 1853
Db   2884 GATCACCACCAGAGCATCGTCGTGTCTGGGGCAAGACTCCCCAAGTTCCAAGCTGCCCATACA 2943
Qy   1854 GAAGGAGACCTGGGAGACCTGGTGGACCGACTACTACGGCAAGGGCAAGAGGGCCACCAGTG 1913
Db   2944 GAAGGAGAGCATGGGAGGACCTGGTGGACGACTACTGGCAAGGGCCACCACCTGGATCCTGAGTG 3003
Qy   1914 GGAGTTCGTGAACACCCCTCCTGGTGTGTAGTCTGGTGTGTATCAGCTGAGGAGGAAGGAACCCAT 1973
Db   3004 GGAGTTCGTGAACACCCCTCCTTGGTGTGAAACTGTGGAGCTGAGGAGGAGAACCCAT 3063

Qy   1974 CATCGGCGGCCGCGAGACCTTCTACGTGGACGGCCGCGGCCAACCCGAGACCAAGATCGGCAA 2033
Db   3064 CGTGGGAGGCAGAGACCTTCTACGTGGATGGGGCCGGCCAGCCAACCAAGGCCAAGAACCAAGCTGGGCAA 3123
Qy   2034 GGCGGCTACGTGACCGACCGGGGCCCGGCAGAGATCTGAGCGTGGACGAGACCACCAA 2093
Db   3124 GGCAGGCTACGTGACCGACCAACCAGGACGAGAACGATCTGACCCTGACTGACCACCACCAA 3183
Qy   2094 CCAGAAGACCGAGCTGCAGGCCATCCAGCTGCCCCTGCAGGACGAGCAGCCGCCAGCGAGGTGAA 2153
Db   3184 CCAGAAGACTGAGCTGCAGGCCAATCCATTAGTCTGCAAGACAGCAGGACTGGAAGTGAA 3243
Qy   2154 CATCGTGACCGACGAGCCAGTAGCGGCCCTGGCACTATCCAGGGCCATCAGCTGGAGCGAGAGGA 2213
Db   3244 CATCGTGACAGACTCACAGTACG-CATGGGCATCATCATCCAAGCACAACCAGATCCGA 3302
Qy   2214 GAGCGAGCTGGTGAACCAGACCATCATGGCAGCTGATCAACGATCATGATCAAGAAGGAGAAGGAGGTGTACCTGAG 2273
Db   3303 GTCAGAGCTGGTGAACCAGCATCATCGAGCAGCTGATCATCGAGAAGGAAGTGTACCTGGC 3362
Qy   2274 CTGGGTGCCCGGCCCACCAAGAGGGCCATCCGGGGCCAAGGAGATCGACAGCGAGATTGAACAAGATGATCATGCAT 2333
Db   3363 ATGGGTACCAGCACCACAAGGAATTGAGGAGAATGAACAAGTAGATAAATTAGTCAGTGC 3422
Qy   2334 GGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGAT 2370
Db   3423 TGGGATCCGGAAGTGCTGTTCCCTGGACGGGGATCCGAT 3459

RESULT 11
US-09-184-418C-8
; Sequence 8, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
;   APPLICANT: Hahn, Beatrice
;   APPLICANT: Gao, Feng
;   APPLICANT: Shaw, George
; TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
; FILE REFERENCE: D6287
; CURRENT APPLICATION NUMBER: US/09/184,418C
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 8
; LENGTH: 9010
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
;   OTHER INFORMATION: isolate=96ZM651; 137..1621:"gag"; 1426..4425:"pol";
;   OTHER INFORMATION: 4370..4948:"vif"; 4888..5178:"vpr";
;   OTHER INFORMATION: 5159..5373-7734..7824:"tat"; 5298..5373-7734..7981:"rev";
;   OTHER INFORMATION: 5387..5647:"vpu"; 5565..8171:"env"; 8173..8793:"nef"
US-09-184-418C-8

Query Match   49.7%; Score 1227.2; DB 4; Length 9010;
Best Local Similarity   70.3%; Pred. No. 2e-194;
Matches 1657; Conservative 0; Mismatches 698; Indels 1; Gaps 1;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCCACATCCTGATGCAGCCGGCAGCAACTTCA 73
Db   1221 TGGCTGAGGCAATGAGCCAAGCCCAAAACAAATAGTGTAAACATACTGACAGAAAAGCAATTTA 1280
Qy   74 AGGGCCCCAAGCGCGATCATCCAAGTGCTTCAACTGCGGCAAGGCAGCACATCGCCCGCA 133
Db   1281 AAGGAAATAAAAGAATGGTTAAATGTTTAACTGTGGCAAGGAAGGGCACATAGCCAGAA 1340
Qy   134 ACTGCGCCGCCCCCCGCCCCCCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
Db   1341 ATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGAAAATGGACACCAAATGA 1400
Qy   194 AGGACTGCACCGAGCCCAGGCCAACTTCTTCCGCGGGAGAACTTGGCCTTCCCCCAGGGCCA 253

```
Db   1401 AAGACTGTACTGAGAGGCAGGCTAA-TTTTTAGGGAAAAATTTGGCCTTCCCACAAGGGA 1459          Qy   1334 ACCAGAAGGAGCGCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCG 1393
Qy    254 AGGCCCCGCGAGTTCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGAGCTGC 313           Db   2540 ATCAGAAAGAACCCCCATTTCTTTGGATGGGTATGAACTCCTGACAAATGGACAG 2599
Db   1460 AGGCCAGGGAATTTCCTTCAGAACACAGGCCAGAGCCAACAGCCCCACCAGCAGAGAGCTTC 1519

Qy    314 AGGTGCGGCGGCACAACCCCCAGCGAGGCCGGCGCCGAGCGCAGGGCACCCTGACT 373           Qy   1394 TGCAGCCCATCGAGCTGCCCGAGAAGAGGAGCTGGACCGTGACGACATCCAGAAGCTGG 1453
Db   1520 AGGTTCGAGGAGCAGACAACCCCCTCCGAAGCAGGAGTCGAAAGACAGGGAGCCTTAACT 1579        Db   2600 TACAGCCTATACAGTCTGGCAGAGAAAAAGATAGTCGACTGTTAATGATATACAGAAGTTAG 2659

Qy    374 TCCCCCAGATCACCCTGGTGCAGCGCCCCCCTGGTGAGCATCAAGGTGGGCGGCAGATCA 433          Qy   1454 TGGGCAAGGTGAACTGGGCCAGCAAGGCCCTGACCAGCTGCGCCAGTGCA 1513
Db   1580 TCCCTCAAATCACTCTTTGGCAGCGACCCCCTGTCTCAATAAAGGTAGGGGCCAAATAA 1639        Db   2660 TGGGAAAATTAAACTGGGCAAGTGCAGGATTTACGGCAGGATTAAAGTAAGCAACTTTGTA 2719

Qy    434 AGGAGGCCCTGCCTGGACACCGGCCGACGACACCGTGCTGCTGGAGGAGATGAGCCTGCCCG 493          Qy   1514 AGCTGCTGCGGCGGCAAGGCCCAGGCCCTGACCGACATCGTGCCCCTGACCGAGGCGGCAGC 1573
Db   1640 AGGAGGCTCTCTTAGACACGGGACGGCAGGTGATCACAGTATTAGAGAAATAAATTTGCCAG 1699        Db   2720 AACTCCTTAGGGGACCCAAAGCACTAACAGTAGTACACTGAGAAGAGCAGAAT 2779

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGCATCGGCGCTTCATCAAGGTGCGCCAGTACG 553           Qy   1574 TGGAGCTGCCGAGAACCGCGAGATCCTGCGGCCGGTGCTACGAGCGTGTACGACGACC 1633
Db   1700 GCAAATGGAAACCAAAAATGATGAGAGCAATTGGAGGCTTTATCGAAGTAAGACAATATG 1759        Db   2780 TAGAATTGGCAGAGAACAAGGAAATTTTAAAAGAACCAGTACATGGGGTATATTATGACC 2839

Qy    554 ACCAGATCCTGATCGAGATCTTGCGCAGGAGGCCATCGGCACCGTGCTGATCGGCCCCA 613           Qy   1634 CCAGCAAGGACCTGCTGCCGAGATCCAGAAGCAGGGCCACGACCAGTCGACCTACCAGA 1693
Db   1760 ATCAAAATACCTATGGAAATTTGTGGAAAAAGCCTAGTAGTTACAGTATTAGTAGGACCTA 1819        Db   2840 CATCAAAAGACTTGATAGCTGAAATACAGACGAAAACACAGGGCATGGACCAATGACATATCAAA 2899

Qy    614 CCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGCGCTGCCACCCTGACTTCC 673           Qy   1694 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGGCCAGATGCGCACCGCCC 1753
Db   1820 CACCTGTGTCAACATAATTGGAAGAAAAATATGTGACTCAGCTTGGATGCACACTAAATTTTC 1879       Db   2900 TTTACCAGGAACCATTCAAAAATCTGAAACACGGGAAGTATGCAAAAATGAGGACAGCGCC 2959

Qy    674 CCATCACGCCCATCGAGACCTGCCCGTGAAGCTGAAGCCCATGGACGGCCCCAAGG 733            Qy   1754 ACACCAACACGTGAAGCAGCTGACCGAGATCGACTGCCATCGAGAGATCATCG 1813
Db   1880 CAATTAGTCCTTATTGAACTGACCAGCTAAATTAAAGGATAATGGCCGAAATGGCCCAAGG 1939       Db   2960 ACACTAATGATGTAAAAACAGTTAACAGAGAGCAGTGCAAAAAAATAGCCCTGGAGAGCATAG 3019

Qy    734 TGAAGCAGTGCCCCTGACCGGAGAAGAAGATCAAGGCCCTGACCCTGCGAGGAGA 793            Qy   1814 TGATCTGGGGCAAGACCCCCAAGTTCGGCCTGCCCATCCAGAGGACCTGGAGACCT 1873
Db   1940 TTAAAACAATGGCCATTGACAGAGAGAAATAAAAGCTTTAACAGCAATTTGTGAAGAAA 1999       Db   3020 TAATATGGGGAAAGATTCCTAAATTAGACTACCATCCAAAAGAAACATGGGAAACAT 3079

Qy    794 TGGAAGGAGGGCAAGACATCACCAAGATCGGCCCCGAGAACCCTACACACCCCCGTGT 853           Qy   1874 GGTGACCGACTACTGGCGAGCAGCGCCACCTGGAGGAGTTCGTGAACACCCCCC 1933
Db   2000 TGGAGAAGGAAGGAAAAAAATTACAAAAAATTGGGCCTGAAATCCATTCCAGTAT 2059           Db   3080 GGTGGACAGACTATTGGCAAGCCACCTGATTTCCTGAGCTGGGGAGTTTGTTAATACCCCCTC 3139

Qy    854 TCGCCATCAAGAAGAGGACAGCACACGTGCTGCGAAGCTGGTGAGCTTCCGCGAGCTGA 913           Qy   1934 CCCTGGTGAAGCTGTGGTACCAGGCCCATCATCGCGCCGAGACCTTCT 1993
Db   2060 TTGCCATAAAAAAAGAGACGACACTACTAAGCTGCCGAAAAATTAGTAGATTTCAGGGAACTCA 2119       Db   3140 TCTTAGTAAAATTATGGTACCAGCTGAGAAAGAACCCATAGTAGGAGGACAGAAACCTTCT 3199

Qy    914 ACAAGGCCACCCAGGACTTCTCGGAGTCGCAGCTGCAGCTCGGCATCCGGCTGGGCATCCCCCGCCGGCTGA 973           Qy   1994 ACGTGGACCGGCGCGCCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACC 2053
Db   2120 ATAAAAAGAACTCAAGACTTTTTGGGAAGTTCAATTGGGAATAACCACCCCAGCAGGGTTAA 2179       Db   3200 ATGTAGATGGAGCAGCAGCCAATAGGGAAACTAAATTAGGAAAAGCAGGGTATATTACTGACA 3259

Qy    974 AGAAGAAGAAGAGCTGACCGTGCTGCTGCGAGCTGAGTGCTGCGAGCGTGAGCGGCGAGGATCGTGACCTGGGG 1033           Qy   2054 GGGGCCGGCAGAGATCGTGAGCCTGACCGAGCTGCCGAGCTGCGAGTCGCAGG 2113
Db   2180 AAAGAAAAAAATCAGTGACAGTAGGAAGGCAAAAAATCTTAGAGCCTTTACTAGCGGATGCAA 2239       Db   3260 GAGGAAGGCAAAAAATTGTTACTCTAAACTGAAACAAAATCAGAAGACTGAATTACAAG 3319

Qy   1034 ACGAGGACTTCCGCAAGCTACACAACGGCCTTCGAGATCCCCAGCATCAACAACGAGACCCCG 1093           Qy   2114 CCATCCAGTCTGGCCTGCCCTGACGACAGGCGGCAGCATCGTGACCGACGCAGCCAGT 2173
Db   2240 ATGAAAAGCTTCAGGAATAATACTGCATTCACCATACCTAGTAACAATGAAAACACCAG 2299       Db   3320 CAATTTACCTAGCTTTGCAAGATTCAGGATCAGAAGTAAACATAGTAACTGACTCACAGT 3379

Qy   1094 GCATCCGTCACCAGTACAACGTGCTGCCCCAGCTGGAAGGGCAGCCCAGCATCTTCC 1153           Qy   2174 ACGGCCCTGGGCATCATCCAGACACCGGCTGGTGCGTCGAGCTGGTGAACCAGA 2233
Db   2300 GGATTAGTAGGAGTCAATATCAATAAATGTGCTTCCACAGGGATGCAAAGGATCACCAGCAATATTCC 2359       Db   3380 ATGGCTTAGGAATCATTCAAGCACATCCAGATAAGAGTGAATCAGAGTTAGTCAACCAGA 3439

Qy   1154 AGGAGCAGCATGAACCTGGAGCTGCCGGAGCGTGTGAGGCTGCACTCTGGAGGGCCGACATCGTGAGATCT 1213           Qy   2234 TCATCGAGCAGCTGATCAACAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCACAAGG 2293
Db   2360 AGAGTAGCATGACATGAAAATCTTAGGGCACAAAATCCAGACACATAGTCATCT 2419       Db   3440 TAATAGAACAATTAATAAAGAAGGAAGGAGAAAGGGTCTACCTGTCATGGGTACCAGCCACATAAAG 3499

Qy   1214 ACCAGTACACACGAGCACCTGTACGTGGGGCAGCAGCCGGCAGCCACCTGGAGATCGGCGCGCCA 1273           Qy   2294 GCATCCGGCGCGGCAACGAGCAGAGATCGACAAGCTGTGGAGCAAGGGCATCCCCAAGGTGCTGT 2353
Db   2420 ATCAATATATGGATGACCTGTATGTGGGGATCTGACTTAGAAATAGGGCAACATAGAGCAA 2479       Db   3500 GAATTGGAGTGAATGAACAGGTAGATAAATTAGTAAGCAGGAATCAGGAAAGTGCTGT 3559

Qy   1274 AGATCGAGGAGCTGCGCCAAGCACCTGCTGCTGGGCTTCACCCCCCGACAAGAAGC 1333           Qy   2354 TCCTGGACGCGCATCGA 2369
Db   2480 AAATAGAGAGAGTTAAGAGAACACATCATTATTAAAGTGGGGGATTTACCACCAGACAAAAC 2539       Db   3560 TTCTAGATGGAATAGA 3575
```

```
RESULT 12
US-09-184-418C-9
; Sequence 9, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
;   APPLICANT: Hahn, Beatrice
;   APPLICANT: Gao, Feng
;   APPLICANT: Shaw, George
;   TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
;   TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
;   FILE REFERENCE: D6287
;   CURRENT APPLICATION NUMBER: US/09/184,418C
;   CURRENT FILING DATE: 1999-11-02
;   NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 9
;   LENGTH: 8972
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
;   FEATURE:
;   OTHER INFORMATION: isolate=96ZM751.3; 137.1632:gag; 1419.4435:pol;
;   OTHER INFORMATION: 4380.4958:vif; 4898.5188:vpr; 5169.7814:tat;
;   OTHER INFORMATION: 5308.7938:rev; 5407.5667:vpu; 5585.8128:env;
;   OTHER INFORMATION: 8130.8753:nef
US-09-184-418C-9

Query Match            48.4%;  Score 1196.2;  DB 4;  Length 8972;
Best Local Similarity  70.0%;  Pred. No. 2.6e-189;
Matches 1662;  Conservative    0;  Mismatches 693;  Indels 20;  Gaps  3;

Qy     14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGGCAGCAACTTCA     73
          ||| ||| || ||||||||| |||||| |   ||||  |  ||||||| ||||| || | 
Db   1214 TGGCTGAAGCAATGAGCCAAGCCACCTGGGATTAGAAACAAATGATGATGCAGAAAGCAATTTTA  1273

Qy     74 AGGGCCCAAGCGCATCATCAAGTGCTTCAAACTGCTTGGAGGAGGGCCCCATCCGCCCGCA    133
          | |||||  | | ||||| |||||||||||||| ||  |||| |  |||| ||| |||| |
Db   1274 AAGGCCCTAAAAGATTGTTAAAATGTTTCAAACTGTGGCAAGGAAGGGCACATAGCCAGA    1333

Qy    134 ACTGCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGCACCAGATGA       193
          | ||| | |    ||| |||||||||| ||||| ||| | ||  | |||||||||||
Db   1334 ATTGCAGGGGCCTCCTGGGAAAAAGGGCTGTTGGAAATGTGAAAAGGACAACCAAATGA   1393

Qy    194 AGGACTGCACCGAGGCGCCAGGCCAGCCAACTTCTTCCGCGGAGGACCTGCCTTCCCCCAGGGCA   253
          |||| ||| || || | ||||| |||| |||||| |||||| |   | |||||||||| |||| 
Db   1394 AAGACTGTACTGAGAGACAGGCCAGCAGGTGCTAATTTTTTAGGGGAAAATTGGCCTTCCAGAAGGGGG  1452

Qy    254 AGGGCCCCGAGTTCCCAGCGGAGCAG----------AACCGGCGCCAACAGCC         295
          |||| || |||| || || || ||||          ||||| | || ||||||
Db   1453 AGGCCCGGGGAACTTCCTTCAGAACAGACCAGAGCCAACAGCCCCACCAGCTCCACCAGCC  1512

Qy    296 CCACCAGCCGCCGGAGTCGCAGGTGCAGGTGCGGGGCGACAACCCCCCCGCCCCCCCC     355
          |||||||| || |  ||  |||| ||||  ||  |||||||||| || | |   |||| 
Db   1513 CCACCAGCAGAGAGCTTCAGGTTCAGTTCGAGCCACCCCCCCCGAGGCAGCAGGAGCAGAAA  1572

Qy    356 GCCAGGGCACCCTGACTTCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCA     415
          |||||| |||| |||  |  | ||| |||||  |||||  ||| |||  |||||  ||
Db   1573 GACAAGGAACCCTTAACTGCCCTCAAATCACTCTTTGGCAGCGACCCCTTGTCTCAATAA  1632

Qy    416 AGGTGGGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCACCGTGCTGG          475
          || | ||| |||||| |||||| | |||| |||| ||||| | |||||||
Db   1633 AAGTAGGGGGGCCAGATAAAAGAGGCTCTCTTAGACACAGGAGCAGATGATACAGTATTAG  1692

Qy    476 AGGAGATGAGCCAGGCCAGTGAAGCTGAAGCCAAGAATGATCGGGGCATCGGGCGGCTTCA     535
          || |||||| |||  ||||||| |||  || |||||| ||   |||||| | | ||  ||
Db   1693 AAGAAATAAATTTGCCAGGAAAATGGAAACCAAAAATGATAGGGGAATTGGAGGTTTTA  1752

Qy    536 TCAAGTGCTGCGCCAGTACGACCAGATCAAAATCGGGGCCATCGGCA         595
          || |||||| |   ||  || |||||| |||| | |   |||||| |
Db   1753 TCAAAGTAAGACAGTATGATCAGATACCCATAGAAATCTGTGGACATAAAGGCTATAGGTA  1812

Qy    596 CCGTTGCTGATCGGCGGCCCCACCCCCGTGAACATCATCGGGCGCAACATGCTGACCCTGG     655
          ||| | |||||| | || |||||||| |||  ||| | || |  ||||| |||| |||| |
Db   1813 CAGTATTAGTAGGACCTACACCTGTCAACATAATTGGGAGAAATCTGAAAACAGAAGTATG  1872

Qy    656 GCTGCACCCTGACTTCCCATCAGCCCATGGACCCGTGCCGTGAAGCTGAAGCCCG       715
          ||||||| || |||||||| | || | | |  ||| |||||||||| || || ||
Db   1873 GCTGCACAGTAAATTTCCAATTAGTCCTATTGAAACTGTACCAGTAAAATTAAAGCCAG   1932

Qy    716 GCATGGACGGGCCCCAAGTGAGCGAGTCAGTGGCCCTGCCCTGACCGAGGAGATCAAGGCCCTGA    775
          ||||| | || ||  ||||| | | | |||| ||||| |||| | | |||| | || ||| ||| 
Db   1933 GAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGAAGAGAAAATAAAAGCATTAA   1992

Qy    776 CCGGCCATCTGCGAGGAGATGAGAAAGGGCAAGATCACCACCAGATCGGCGGCCCCGAGAACC    835
          || ||||| || || | |||| |||| ||||  |||| |  ||||||| ||||| |||||| |
Db   1993 CAGCAATTGTGAAGATGGAAATGGAAGGAAAAATTACAAAAATTGGCCCTGAGAATC  2052

Qy    836 CCTACAACACCCCGTGTTCGCCATCAACAAGAGGACAGCAGCACCAAGTGGCGCAAGCTGG    895
          |||| || | |  || |  | ||| || |||||||| || || |||  ||||| ||||| |
Db   2053 CATATAACACTCCAGTATTTGCCATAAAGAAAAAAGAAAGTACTAAGTGGAGAAAAATTAG  2112

Qy    896 TGGACTTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCCATCC    955
          ||||||| || ||| | |||| | ||| || ||| ||| || | || ||||| | |  |
Db   2113 TAGATTTCAGGGAACTCAATAAAAGAACTCAGGACTTTTGGGAAGTTCAATTAGGAATAC  2172

Qy    956 CCCACCCCGCCGGCTGAAGTGAGAGGCTGTGGACCGTGCTGGACGTGGGCGACGCCT     1015
          || ||||||| ||| |||| ||  || || || || || || || | | ||  |||
Db   2173 CACACCCAGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGGGATGCAT   2232

Qy   1016 ACTTCAGCGTGCCCCTGGACGAGGACTTCCGGAAGTACACCGCCTTCACCATCCCCAGCA    1075
          || |||  ||||| |||||| |||| || || |   | | | | | || || |||| | |
Db   2233 ATTTTTCAGTTCCTTTAGATGAAGGCTTCAGGAAATATACTGATTCACCATACTGTAGTA  2292

Qy   1076 TCAACAAGAGACACCCCCGGCATCCTGATGCAGGCCACCTCGATGGTGCTGTCCCCAGGGGTGGAAGG    1135
          || || |||| ||||||  || |||| | |||| | | || || |  | | || | |  |||||| |
Db   2293 TAAACAATGAAACACCAGGGATTAGATATCAATATAATGTGCTTCCACAGGGATGGAAAG  2352

Qy   1136 GCAGCCCCAGCTCTTCCAGAGCAGCATGACTGTCAAGATCCTCAAGATCGATAACTTTAGAGACACCTTTAGGACACAAA   1195
          || | || |||| ||||| || ||||| |  | || | | ||||||| || | |    |
Db   2353 GATCACCATCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAGCACAAAA  2412

Qy   1196 ACCCCGAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGA    1255
          || |||| || | | || |||| || | |||  | ||| | |||||  ||| | |||||
Db   2413 ATCCAGAAATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGATTTAGAAA  2472

Qy   1256 TCGGCCAGCACCGGCGCCAAGATGAGCGGGCATCGTCGCAAGCACCACCGCTGCTGGGCTTCA    1315
          | |||| |||  |  ||| ||| |||| ||| || |  |||||| |  || | ||||||| |
Db   2473 TAGGGCAACATAGAGCAAAAATAGAGGAGTTAAGAGAACATCTATTGAAGTGGGGATTTA   2532

Qy   1316 CCACCCCCGACCCCTTCGTCTTCTTTGGATGGGGTACGAGCTGC     1375
          || || ||||| || |  | ||||| || ||| || |  ||| |
Db   2533 CTACACCAGACAAAAAACATCAGAAAGAACCTCCATTTCTTTGGATGGGGTATGAACTCC   2592

Qy   1376 ACCCCGACAAGTGGACCGTGCCCAGCGAGGTGCCGCGGCCGAGAGCTGACCGTGA     1435
          || |||| |||||  | |||||| |||| | |||  | || ||||  |   | ||
Db   2593 ATCCTGACAAATGGACAGTACAGCCTATAAAGCTGCCAGAAAAGGAGAGCTGGACTGTCA   2652

Qy   1436 ACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCA    1495
          | || ||| ||| | || | |  | |||| |  | |||||||| ||   ||||  || |
Db   2653 ATGATATACAGAAGTTAGTGGGAAAAT-AAACT-GGCAAGTCAGATTTACCCAGGGATTA   2711

Qy   1496 AGGTGGCGCCAGTGTGCAAGCTGCTGTGCGCGGCGCCCAAGGCCCGACATCGTGCCCC    1555
          || | |  |||| | |  ||    |||  | || ||| ||||  | |  | | |||| |
Db   2712 AAGTAAGGCAACTGTGTAAACTCCTTAGGGGAGCCAAAGCACTAACAGACATAGTACCAT  2771

Qy   1556 TGACCGAGGAGCCAGCTGGAGCTGGAGAATTAGAATTGGGGCGGCATCGGGGCATCGGAGCCCCTGC    1615
          ||| |||||| || || || | | ||||  |||| || || | || | | | ||||  |  |||| |
Db   2772 TGACTGAAGCAGCAGAGTTAGAATTGGCAGAAAACAGGGAATTGGAAAATACAGTATTTTA   2831

Qy   1616 ACGGCGTGTACTACGACCCAGATCTGATCGAGCTGTGCGGGCAAGAGGGCCATCGGCA     1675
          | || || | || || ||| | | | | |  || |  |  || ||| || |  |   |
Db   2832 ATGGAGTATTATGCAGACCCATCAAAAGACTTAATAGCAGAAATACAGAAACAAGGGCATG  2891

Qy   1676 ACCAGTGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACTGCTGACCCAGCTGG    1735
          | ||||||||| || |  | |||| |  |||| |  ||  | |||||   | | |||| | ||
Db   2892 ACCAATGGACAGATCAGTATCAAGTGCTGCCAACAATTGTTGACCCAGCTTG  2951
```

Matches 1337; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

```
Qy   790  GAGATGGAGAGGAGGGCAAGATCACCCAGAATCGGCCCGAGAACCCTACACACACCCC   849
Db     7  GAGATGGAGAGGAGGGAGGAGATCAGCAGATCGGCCTGAGAACCCTACAACACTCCA    66

Qy   850  GTGTTCGCCATCAAGAAGAAGCAGCAGCAAGTGGCCAAGCTTGTGACTTCCGCGAG     909
Db    67  GTCTTCGCAATCAAGAAGAAGCAGTACCAAGCTGTGAGAAAGCTGTGGACTTCAGAGAG   126

Qy   910  CTGACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGCATCCCCAGCCCGGCGGC     969
Db   127  CTGACAACAGCAACTCAGGACTTCTGGGAAGTTCAGCTGGCATCCCCAGTCGAGTGGG    186

Qy   970  CTGAAGAGAAGAAGCGTGACCGTGCTGGGCGACGCGCTACTTCAGCGTGCCC       1029
Db   187  TTGAAGAGAAGAAGTCAGTGACAGTGCTGGATGTGGTGGATCCAGTTCTCCGTTCCC    246

Qy  1030  CTGGACGAGACTTCCGCAGTACACCGCCTTCACCATCCCAGCATCAACAACGAGACC   1089
Db   247  TTGGACGAGACTTCAGGAAGTACAGATGTACAACTGCCTTCACGATACCTAGCATCAACGAGACA  306

Qy  1090  CCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATC 1149
Db   307  CCAGGGCATCCGCTACCAGTACAACGTGCTGCCACAGGGATGGAAGGGATCACCAGCATC   366

Qy  1150  TTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCGCAGCGAACCCGAGATCGTG 1209
Db   367  TTTCAAAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGAAGCAAAACCCAGAGACATCGTG  426

Qy  1210  ATCTACCAGTACATGGACGATCTGTACGTGGGCAGCGACCTGGAGATCGGCACCACCGC   1269
Db   427  ATCTATCAGTACATGGACGATCTTCACGTAGGAAGTGACCTGGAGATCGGCAGCACAGG    486

Qy  1270  GCCAAGATCGAGGAGGCTGCGCAAGCACCTCGCTGCCGTGGGCTTCACCCCGACAAG    1329
Db   487  ACCAAGATCGAGGAGGCTGAGACTGTTGAGGTGATGGGTGACCACCAGGACAAGACAAG   546

Qy  1330  AAGCACCAGAGAGGAGCCCCCTTCGTGTGAGGAAGGAGAGCTGACCGTGAACGACATCCAGAAG 1389
Db   547  AAGCACCAGAGAGGAACCTCCCTTCGTGCTGCTGTGGATGGGTACGGACTGCATCCTGACAAGTGG 606

Qy  1390  ACCGTGGGCAAGCTGAACTGGGCCAGCAGCATCTACCCCGGCATCAAGGTGCGCCCAGCTG 1449
Db   607  ACAGTGCAGCCCATCGAGCTGCGCTGCCTGAGCAGCATCACCCAGGCCATCGTGAACGACAATACAGAG 666

Qy  1450  CTGGTGGCAAGCTGAACTGGGCCAGCAAGGCCTGACCGCATCGTGCCCCTGACCAAGGT   1509
Db   667  CTCGTGGGCAAGTTGAACTGGGCAAGCCAGGCCTGAGAACCTCACCCAGGCCATCAAAGTTAGGCAGCTG 726

Qy  1510  TGCAAGCTGCTGCGCGGCGCCAAGGCCTGAACTGGGCCAGCATCGTGCCCCTGACCGAGGAGGCC 1569
Db   727  TGCAAGCTGCTTCGAGGAACCAAGGCACTGACAGAAGTGATCCCACTGACAGAGGAAGCA 786

Qy  1570  GAGCTGGAGCTGGCCGAGACCGAGACTGCTGCGCGAGCCCGTGCACGCGCGTGTACTAC 1629
Db   787  GAGCTAGAACTGGCAGAGAACCGAGACTGCTGCGAAGGAGCCAGTGCAATGGAGTGTACTAC 846

Qy  1630  GACCCCAGCAAGGACGACCTGGTGGCCGAGATCCAGAAGCAGGCCAGGGCCAAGTACGGCCACC 1689
Db   847  GACCCAAGCAAGGACGACCTGATCGCAGATCGCAGAATCCAGAAGACCAGAGGGCAAGGGCCAATGGACCTAC 906

Qy  1690  CAGATCTACCAGGAGCCCTTCAAGAACTGGGAGAGCATGGCCAGCGACTACGGCCAGGAAGAGC 1749
Db   907  CAAATCTACCAGGAGCCCTTCAAGAACTGGGAGAGCATGGCCAGCGACTACGGCCAGGAGAGGGGT 966

Qy  1750  GCCCACACCAACGACGTGAAGCAGTGACCGAGAAGATCGTGCAGAAGATCGGCGATGGAGAGC 1809
Db   967  GCCCACACCAACGATGTGAAGCAGTGACAGGCAGTGCAGAGATCACCACAGAGAGC 1026

Qy  1810  ATCGTGATCTGGGGCAAGACCCCCAAGTTCGCGCTTGCCCATCCAGAAGGAGACTTGGGAG 1869
Db  1027  ATCGTGATCTGGGGCAAGACTCCCCAGTTCAAGCTGCTGCCCATACAGAGGAGGAGGACATGGGAG 1086
```

```
Qy  1736  CCAAGATCGCGCACCGCCCACCAACCAAGACGAGCTGAAGCAGCTGACCGACGCCGTGCAGAAGA 1795
Db  2952  CAAAAATGAGGAGCTGCCCACTAATGATGTGTAAAACAGTTAACAGAGAGGCGGTGCAAAAAA   3011

Qy  1796  TCGCCATGGAGAGCATCGTGACTGGGCAAGACCCCCAAGTTCCGGCCTGCCCATCCAGA   1855
Db  3012  TAGCCATGGAAGCATAGTAATATGGGAAAGATTCCTAAATTTAGGCTACCATTCAAA    3071

Qy  1856  AGGAGACCTGGGAGACCTGTGACGACTACTGGCAGGCCACCTGAGATCCCCAGTGGG   1915
Db  3072  AAGAAAACATGGGAGACATGGTGACAGACAGAGCTATTGGCAAGCCACCTGAGATTCCTGAGTGGG 3131

Qy  1916  AGTTCGTGAACACCCCCCCTGGTGAAGCTGGTACCAGTGGGAAGGAGCGCCATCA    1975
Db  3132  AGTTTGTTAATACTCCCCCCCCAGTAAAATATGGTACCAGTGGAGAAAGAAAGACCCCATAG 3191

Qy  1976  TCGCGCCGCCACCTTCCACGTGGACGCGCGCCAACCGCGCGAAGATCGGCGCAGG    2035
Db  3192  TCGGCGCCAGACCTTCTACGTGGACTGGAGCAGCAAGGGCCAATAGGAGAAACTAAAATAGGAGAAAG 3251

Qy  2036  CCCGGCTACGTGACCGACCGGGGGCGGCAGAAGATGTGAGCCTGACGGCAGCCAACC 2095
Db  3252  CAGGGTATGTTACTGACAGAGGAAAGGAAGGAAAAAATTGTTACTCTAACTGAAACAACAAATC 3311

Qy  2096  AGAAGACCGAGCTGCAGGCCATCCAGTTGGCCCTGCAGGACAGCGCCGCGAGGTGAACA 2155
Db  3312  AAAAGACTGAATTACAAGCAATTCAGTTAGCTTTGCAGGATTCAGGATCAGGAAGTAAAACA  3371

Qy  2156  TCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGCGGCCCCAGCGAGAGCAGGTAGAAT 2215
Db  3372  TAGTAAACAGACTCACAGTATATGCATTAGGAATCATCATCCAGCCAAGCACACAACCAGATAAGCAGTGAAT 3431

Qy  2216  GCGAGAGCTGGTGAACCAGATCATCGAGCTGATCATCGAGAAGAGAAGGAGTGTACCTGAGCT 2275
Db  3432  CAGAATTAGTCAATCAAATAATGAACAAGTTGTATAAAAAGGAAAGGGTTTTACCTGGTCAT   3491

Qy  2276  GGGTGCCGCCCCACAAAGGCGCATCGTGGGCGGCAACGAGCAGCAGATCGACGACAAGCTGTGAGCCAAGG 2335
Db  3492  GGGTACCAGCACACAAACGAAGTTGGAGGAAAATGAACAAGTAGATAAATTGGTAAGTAGTG 3551

Qy  2336  GCATCCGCAAGGTGCTGTTCCTGGACGCGGCATCGAT                        2370
Db  3552  GAATCAGGAAAGTGCTGTTTCTAGATGGAATAGAT                         3586
```

RESULT 13

US-09-872-733A-3

```
; Sequence 3, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:  Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Pol gene
```

US-09-872-733A-3

Query Match 48.2%; Score 1190.6; DB 4; Length 2467;
Best Local Similarity 84.6%; Pred. No. 1.9e-188;

```
Qy   1305 TGTGGCAAGGAGGGCACATAGCCAGAAATTGCAGGGCGCCCCTAGAAAAAAGAGGCTGTTGG 1364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 AAGTGCGGCAAGGAGGGGCCACCAGATGAAGGACTGGCACGGAGCGCCAGGCGCCAACTTCTTC  225

Db   1870 ACTTGGTGGACCGACTACTGGCCAGGCCACCTGATCCCCAGTGGGAGTTCGTGAAACACC 1929
          |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1087 ACATGGTGGACCGAGTACTGGCCAAGCCACCTGGATCCCTGAGTGGGAGTTCGTGAAACACC 1146

Db   1355 AAATGTGGGCAAGAAGGACACCAAATGAAAGATGAAAGACTGTACTGAGAGGCAGGGTAA-TTTTTT 1423

Qy    226 CGGCGAGGAGACCTGGCCTTCCCCAGGGAGGAGCCCATCATCGGCGCCGAGACC 1989
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1147 CCTCCCTTGGTGAAACTGTGTGTATCAGCTGTATGCCGCCATCCGTGGGACAGAGACC 1206

Qy   1930 CCCCCCTGGTGAAGCTGTGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACC 1989

Qy   1990 TTCTACGTGGACGGGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACC 2049
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1207 TTCTACGTGGATGGGGCAGCAACCAAGGAGACCAAGATCGGCAAGGCCGGCTACGTGACC 1266

Qy    286 GCCAACAGCCCCAGCCGCGAGACTGCAGGTGCGGGCGACAACCCCGAGCGAGGCC  345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1484 GCCAACAGCCCCACCAGCAGAGAGCTTCAGGTTCAAGGAGCAACAACCCCGTCCGAAGCA 1543

Qy   2050 GACCCGGGGCCGCAAAGATCGTGAGCCTGACCGAGACCACCAACCAACCAGAAGACCGAGCTG 2109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1267 AACCGAGGAGGACGACAGAAAGTGGTGACCCGACTGCAGCGCGCGAGTGAACATCGTGACCGTGACACGAC 1326

Qy    346 GGGCGCCGAGCGCCAGGGAACCTGCGAACTTCCCCAGATCACCCTGTGGCAGCGGCCCCTG  405
          ||| | |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1544 GGAGTCGAAAGACAGGGAACCCTAACTTCCCTCAAATCACTCCTTGGCAGCGACCCCTT 1603

Qy   2110 CAGGGCATCCAGCTGCCCTGCAGCAGCGGGCGCAGCGGAGGTGAACATCGTGACCGACAGC 2169
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1327 CAAGCCATCATCTCTGCAAGACAGCAGCGACTGGAAGTGAACATCGTGACAGACTTCA 1386

Qy    406 GTGAGCATCAAGGTGGCGGCCAGGATCAAGAGGACCCTGTGGACACCGGCCGACGAC  465
          ||| |  |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1604 GTCTCAATAAAAGTAGGGGCCAGATAAAAGGAAGTCTCTTAGACACAGGAGCAGATGAT 1663

Qy   2170 CAGTACGCCCTGGGCATCATCCAGGCCCGACAAGAGAGCGAGAGCCGAGCTGGTGAAC 2229
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1387 CAGTACGCACTGGGCATCATCCAAGCACCAACCAGACCAATCCGAGTCAGAGTCGTGAAC 1446

Qy    466 ACCCTGCTGGAGGAGATGAGCCTGCCCGGCCCAAGTGAAGCCAAGAGTGATCGGCGGCATC  525
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1664 ACAGTATTAGAAGAAATAGCTTTGCCAGGAAGTGAAAACCAAAAATGATAGGAGAGAATT 1723

Qy   2230 CAGATCATCGAGCAGCTGATCAAGAAGAGGAAGTGTACCTGAGCTGAGCTGAGCCCACC 2289
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1447 CAGATCATCGAGCAGCTGATCAAGAAGAGGAAAGTGTACCTGGCATGGTACCAGCACAC 1506

Qy    526 GGGCGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGGACCCGGTGAGATCGGCGGCAAGAAG  585
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1724 GGAGGTTTTATCAAAGTAAGACAGTAATGATCAAATACTTATAGAAATTTGTGAAAAAG 1783

Qy   2290 AAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTTGGTGAGACAAGCTCGGCAAGGTG 2349
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1507 AAAGGAATTGGAGGAAATGAACAAGTAGATAAATTAGTACAGTGCTGGGATCCGGAAGGTG 1566

Qy    586 GCCATCGGCACCGTGCTGATCGGCCGTGATCGGCCCGTGAACATCATCGGCGACAACATGCTG  645
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1784 GCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAGATATGTTG 1843

Qy   2350 CTGTTCCTGACGGCATCGAT 2370
          ||||||||||||||||||||
Db   1567 CTGTTCCTGGACGGGGATCGAT 1587

Qy    646 ACCCAGCTGGGCTGCACCTGCAACCTGAACTTCCCCATCAGCCCCATGCAGCCCGTGCCGTGAAG  705
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1844 ACTCAGCTTGGATGCACTTTAAATTTTCCAATTAGCCCCATTAGCCCCATTGAAATTGTAAAA 1903

Qy    706 CTGAAGCCCGGCATGAGCGGCCCCAAGTGAAGCAGTGGCCCTGACCGAGAGAAGATC  765
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1904 TTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAGTGGCCATTGACAGAAGAGAAAATA 1963

Qy    766 AAGGCCCTGACCGCCATCGTGCGAGGAGGATGGAGAATGGAGGAGGGGCAAGATCACCGAGATCGGC  825
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1964 AAAGCATTAACAGAAATTTGTAAAGAAATGGAAAAGGAAGGAAAAATTTACAAAAATTGGG 2023

Qy    826 CCCGAGACCCCTACAACACCCCGTGTTCGCATCAGCGATTTGCCAGCAGCCAAGTGG  885
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2024 CCTGAAAATCCATATAACACTCCAGTATTTGCCATAAAAAGGAAGGACAGTACTAAGTGG 2083

Qy    886 CGCAAGCTGGTGGACTTCCGGCCAGCTGAACAGCGCACCCGCACCCAGGACTTCTGGAGGTGCAG  945
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2084 AGAAAATTAGTAGATTTCAGGGAGCTCAATAAAAGAACTCAAGACTTTTGGGAAGTTCAA 2143

Qy    946 CTGGGCCATCCCCACCCCTGGAGAGGAGCCTGGACCGTGCTGGACGTG 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   2144 TTAGGAATACCACCCCAGTTTAAAAAGAAAAATCAGTGACTACTGGGATGTG 2203

Qy   1006 GGGACGCCTACTTCAGCGTGCCCTGGACGAGGAGGACTTCCCAAGTACACCGGCCTTCACC 1065
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2204 GGGGATGCATATTTTTCAGTTCCTTTAGATGAAGGCTTCGGGAAATATACTGCATTTACCACC 2263

Qy   1066 ATCCCCAGCATCAACAACGAGACCCCCGCATCCGCTACCAGTACACCGTGCTGCCCCAG 1125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2264 ATACCTAGTATAAACAATGAAACACCAGGGATTAGATATCAATATGTGCTTCCACAG 2323

Qy     46 AACATCCTGATGCAGCCGCAGCAACTTCAAGGGCCCAAGCGCCATCATCCAAGTGCTTCAAC  105
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1245 AACATAATGATCCAGAGAGGCAATTTTAAAGCCCTAAAGAATTGTTAAATGCTTCAAC 1304

Qy   1126 GGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAAGATCCTGGAGCCCTTC 1185
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2324 GGATGGAAAGGATCACCAGCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCCCTTT 2383

Qy    106 TGCGGGCAAGGAGGGGCCACATCGCCCGCCCCCGCCAACTGCCCGCGCCAACTCGGGGGGCCCCCCCCGCAAGAGAGGGGGCTGCTGG  165
          ||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2384 AGGGGCACGAAATCCAAAAAATAGTCATCTATCAATATATGGATGACTTGTATGTAGGGTCT 2443
```

```
Query Match             48.2%;  Score 1189;  DB 4;  Length 8959;
Best Local Similarity   70.0%;  Pred. No. 4e-188;
Matches 1627  Conservative  0;  Mismatches 696;  Indels  2;  Gaps 2;
```

```
Qy    2326  GTGAGCAAGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGAT  2370
                                                           
Db    3523  GTAAGTAGTGAATTAGGAGAAGTACTGTTTCTAGATGTGGATAGAT  3567
```

RESULT 15
US-09-184-418C-4
; Sequence 4, Application US/09184418C
; Patent No. 6492110
; GENERAL INFORMATION:
;   APPLICANT: Hahn, Beatrice
;   APPLICANT: Gao, Feng
;   APPLICANT: Shaw, George
;   TITLE OF INVENTION: CLONES AND SEQUENCES FOR NON-SUBTYPE B ISOLATES OF HUMAN
;   TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPE 1
;   FILE REFERENCE: D6287
;   CURRENT APPLICATION NUMBER: US/09/184,418C
;   CURRENT FILING DATE: 1999-11-02
;   NUMBER OF SEQ ID NOS: 112
; SEQ ID NO 4
;   LENGTH: 8992
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
;   FEATURE:
;   OTHER INFORMATION: isolate=92RW009; 139.1624:gag; 1690.4428:pol [N-terminus uncertain
;   OTHER INFORMATION: 4373.4951:vif; 4891.5181:vpr; 5162.7801:tat; 5301.7958:rev;
;   OTHER INFORMATION: 5403.5648:vpu; 5566.8148:env; 8150.8773:nef
US-09-184-418C-4

  Query Match   47.2%; Score 1165.6; DB 4; Length 8992;
  Best Local Similarity   68.9%; Pred. No. 2.9e-184;
  Matches 1627; Conservative   0; Mismatches 729; Indels   4; Gaps   2;

Qy   14 TGGCCGAGGCCATGAGCCAGGCCACCATCCTGATGCAGCGCAGCAACTTCA   73

Db   1221 TGGCTGAGCAATGAGCCAGGCCAAGTACAACAACCAAACATAATGATGCAGAGAGGCAATTTTA 1280

Qy   74 AGGGCCCCAAGGCGCATCATCAATGCTTCAACTGCGGCAAGGAGGGCCACATCCGCCCGCA   133

Db   1281 AGGGCCAGGAACAAGAAGATTATTAAGTGTTTCAACTGTGGCAAAGAAGGACACCTAGCCAGAA 1340

Qy   134 ACTGCGCGCCCCCCCCAAGAGAGGGCTGCTGGAAGTGCGGCAAGGAGTGCGGCACCCAGATGA   193

Db   1341 AATTGCAGGGCCCCTAGAAAAAAGGGCGTTGGAAATGTGGAAAGGACACCAAATGA   1400

Qy   194 AGGACTGCACGCAGGCCAGCCAACTTCTTCCGGCAGGAGACCTGGCCTTCCCCCCAGGGCA   253

Db   1401 AAGACTGCACTGAGAGACAGGCTAA-TTTTTTAGGGAAAATTTGGCCTTCCAACAAGGGG 1459

Qy   254 AGCCCCGGAGTTCCCCGGCGACAACCCCGCCGAGCGAGCGCCGGCGGCCGAGCGGCACCCTGA   311

Db   1460 AGGCCAGGAAATTTCCCCAGCAGCACTGGGCCAACAACAGCCCCACCAGCAGCAGAGAACTTT 1519

Qy   312 -GCAGGTGCGCGGCGACAACCCCGCGGGCGAGGGCCGAGCGGCCGCCAGGCGCACCCTGA   370

Db   1520 GGAATGGGGGAAGCAGAGATAGCCTCTCCTTCTGAAACAGGGAACCAGATCGTGAGCGTGAGCGGG 1579

Qy   371 ACTTCCCCAGATCACCCTGTGCAGGCCCATCACCCTGGTCAGGGCGGCAGGAA   430

Db   1580 ATTTCCCTCAAATCACTCTTTGGCAACGACCCCTTGTCACAGTAATAGGAGGACAGGTTCAGC 1639

Qy   431 TCAAGGAGGCCCTGCTGGACACACCGGCCGACGACACCGTGCTGGAGGGAGATGAGCCTGC   490

Db   1640 TAAGAGAGCTCTTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATAAATTTGC 1699

Qy   491 CCGGCAAGGTGGAAGCCCAAGATGATCCGATCAAGAAGGAGAAGGTG   550

Db   1700 CAGGAAAATGGAAACCAAAAATGATAGGGGAATTGGAGGTTTATCAAGGTAAAACAGT 1759

Qy   551 ACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCCGACAAGGCTG   610

Db   1760 ATGATCAAATACTTATAGAAATTTGTGGAAAAAAGGCTATAGGTACAAGTATTAGTAGGAC 1819
```

```
Qy    1246  GACTCTGGAGATCGGCCAGCACCCGCCCAGAATCGAGGAGCTCGCCAAGCACCTGCTGCGC  1305
                                                                        
Db    2444  GACTTAGAATAGGGCATCATAGAGCAAAAATAGAGGAGTTAAGAGCACATCTATTAAAG  2503

Qy    1306  TGGGGCTTCACCACCCCCGACAAGAAGCACCCAGAAGAGACGCCCCCTTCCTGTGGATGGGC  1365
                                                                        
Db    2504  TGGGGATTCACCACACCAGACATCAGAAGAACATCAGAAAGAACCCCATTTCTTTGGATGGGG  2563

Qy    1366  TACGAGTCTGCACCCGACAAGTGACCGTGCAGCCCATCGAGCTGCCCGAGAGGAGAGAGC  1425
                                                                        
Db    2564  TATGAACTTCCATCCTGACAAATGGACAGTACAGCCTATAAAAGTGCCAGAAAAGGATAGC  2623

Qy    1426  TGGACCGTGAACGACATCCAGAAGCTGTGGGCCAAGCTGAACTGGGCCAGGATCTAC  1485
                                                                        
Db    2624  TGGACTGTCAATGATATACAGAAGTTAGTGGGAAAATTAAACTGGGCAAGTCAGATTTAC  2683

Qy    1486  CCCGGCATCAAGGTGCGCCCAGCTGCTGCAAGCTGCTGCCGCGCGGCCCTGACCGAC  1545
                                                                        
Db    2684  CCAGGGATTAAAGTAAAGCAACTTTGTAAACTCCTTAGGGGCCAAAGCACTAACAGAC  2743

Qy    1546  ATCGTGCCCCTGACCAGTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGC  1605
                                                                        
Db    2744  ATAGTACCACTAACTGAAGAAGCAGAATTAGAATTAGCAGAGAACAGGGAAATTCTAAAA  2803

Qy    1606  GAGCCGGATCAAGGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTCGTGTGGCCGAGATCCAGAAG  1665
                                                                        
Db    2804  GAGCCAGTACATGGAGTATTATGACCCATCAAAAGACTTAATAGCTGAAATACAGAAA  2863

Qy    1666  CAGGGCCCACGACCAGTGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACC  1725
                                                                        
Db    2864  CAGGGGCATGACCAATGGACACATATCAAATTTACCAAGAACCATTCAAAAATCTGAAAACA  2923

Qy    1726  GGCAAGTACGCCAAGATCGCACCGCCACCAACCAACGTGAAGCAGCTGACCGAGGCC  1785
                                                                        
Db    2924  GGGAAGTATGCAAAAATGAGGGGTGCTCACACTAATGATGTAAAAACAGTTAACAGAGGCA  2983

Qy    1786  GTGCAGAAGATCGCCATGGAGGAGATCGTGAGCGAGGCCAAGAGCTGGGCAAGTTCCGCCTG  1845
                                                                        
Db    2984  GTGCAAAAAATAGCCATGGAGAGCATAGTAATAT-GGGAAAAGCCCTAAATTAGGATCTA  3042

Qy    1846  CCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCTACTGGCAGGCCACCTGGATC  1905
                                                                        
Db    3043  CCCATCCAAAAAGAAAGTGGGACACTATTGGCAGGACTATTGGCAGGCCACCTGGATT  3102

Qy    1906  CCCGAGTGGGAGTTCGTGAACACCCCCCCCCCTCGTGGTGGTACCAGCTGGGAGAAG  1965
                                                                        
Db    3103  CCTGATTGGGAGTTTGTTAATACCCCTCCCCCAGTAAATTATGGTACCAGCTAGGAAAAA  3162

Qy    1966  GAGCCCATCATGGGCCGCCAGACCTTCTACGTGACGCCGGCCGCCCAACCGCGAGACCAAG  2025
                                                                        
Db    3163  GAACCCATAGTAGGAGCAGAAACTTTCTATGTAGATGGGGCAGCTAATAGGGAAACTAAA  3222

Qy    2026  ATCGGCCAAGGCCGGCTACGTGACCGAGCCGGGGGGCGCCGAGAGATCGTGAGCGACGGCGAG  2085
                                                                        
Db    3223  GTAGGAAAAGCAGGGTATGTTACTGACAGAGGAAGGCAGAAAAATTGTTTCTTTAACTGAA  3282

Qy    2086  ATCGGCGAGAACCAGAGAGACCGGCTGCAGGACGCCAGGCCATCCGCCGGCGGCAGC  2145
                                                                        
Db    3283  ACAACAAATCAGAAGACTGAATTGCAAGCAATTCAGCTTTGCAAGATTCAGGACA  3342

Qy    2146  GAGGTGAACATCGTGACGACGCACTGGTGACCAGGCCCCAGCCCCGAC  2205
                                                                        
Db    3343  GAAGTAAAACATAGTAACAGACTCACAGTATGCACCAGGAATCATTCAAGCACAACCAGAT  3402

Qy    2206  AAGAGCCGAGAGCGAGCTGGTGAACCAGATCATCGAGCCAGCTGATCAAGAAGGAGAAGGTG  2265
                                                                        
Db    3403  AAAAGTGAATGCAGAGTAGTCAATCAAATAATAGAACAATTAATAAAAAAGGAAAAGAGTC  3462

Qy    2266  TACCTGAGCTGGGTGCCCGCCCACAGGGCATCGGCGACATGGAATTGGAGGGAATGAACAAGCTG  2325
                                                                        
Db    3463  TATCTGTCATGGGTACCAGCACATAAAGGAATTGGAGGAAATGAACAAGCAAGTAGTAGGAC  3522
```

Qy 611 CCACCCCGTGAACATCATCGCCGCAACATGCTGACCCAGCTGGGCTGCACCTGAACT 670
Db 1820 CTACATCTGTCAACATAATATGTTGAAGAAATATGTTGACCCAGATTGGTTGTTACTTTAAACT 1879

Qy 671 TCCCCATCAGCCCCATCGAGACACGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCA 730
Db 1880 TTCCAATTAGTCCTATTGAGACTGTACCAGCTAGCATTAAAGCCAGGAATGGATGGCCCAA 1939

Qy 731 AGGTGAAGCAGTGGCCCTGACCGAGAGAGAGTCAAGGCCCTGACCGCCATCTGCGAGG 790
Db 1940 AGGTTAAACAATGGCCATTGACAGAGAAAAATAAAGCATTAAGAGAAATTTGTACAG 1999

Qy 791 AGATGAGAAGGAGGGCAAGATCACCAAGATCGGCCCGAGAACCCTACAACACCCCCG 850
Db 2000 AAATGGGAAAAGAGGGAAAATTCAAAAATCGGGCTGAAAATCATATAACACTCCAG 2059

Qy 851 TGTTCGGCCATCAAGAAGAAGCAGCACCAAGTGGGCAAGCTGGTGGACTTCCGCGAGC 910
Db 2060 TATTTGCCATAAAAGAAGGACAGTACTAAGTGGAGAAAATTAGTAGATTTCAGGGAAC 2119

Qy 911 TGAACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGCATCCCCCGCGCCGGCC 970
Db 2120 TCAACAAAAGAACTCAAGACTTTTGGGAAGTCCAATTAGGGATACCCACCAGCAGGGT 2179

Qy 971 TGAAGAAGAAGAAGAGCCGTGACCGTGCTGGACGTGGCGACGCCTACTTCAGCGTGCCCC 1030
Db 2180 TAAAGAAGAAAAAATCAGTGACAGTACTGGATGTGGGGATGCATACTTCTCAGTTCCTT 2239

Qy 1031 TGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCC 1090
Db 2240 TAGATGACAGCTTCAGGAAAATATACTGCATTCACCATATCCAGTATAAACAATGAAAACAC 2299

Qy 1091 CCGGCATCCGTCACGTACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCT 1150
Db 2300 CAGGAATTAGGTATCAATATAATGTGCTTCCACAGGATGGAAAGGATCACCAGCAATAT 2359

Qy 1151 TCCAGAGCAGCAAGATCCTCCTGGAGCCCTTCCGGCCCCAACCCCGAGATCGTGA 1210
Db 2360 TCCAAAATAGTAGTGACAAAAATCTTAGAGCCCTTTAGGGCACAAAACCAAGAAATAGTGA 2419

Qy 1211 TCTACCAGTACATGGACGACCTGTACGTGGGGCAGCCACCTGGAGATCGGCCAGCACCGCG 1270
Db 2420 TCTATCAATATATGGATGACATCTGTGTACCTAGGATCTGACTTAGAAATAGGGCAACATAGAG 2479

Qy 1271 CCAAGATCCGAGGAGCTGCGCAAGCACCTGCTGCGTGGGGCTTCACCACCCCGACAAGA 1330
Db 2480 CAAAAATAGAGGAGTTAAGAGAACATCTATTAAAGTGGGGATTTACCACCACCAGACAAGA 2539

Qy 1331 AGCACCGAAGAGGAGCCCCCTTCGTGGATGGGCTACGAGCTGCACCCGGACCAAGTGGA 1390
Db 2540 AACATCAGAAAAGAAACCTCCATTTCTTTGGATGGGTATGGAACTTCATCCTGACCAAATGGA 2599

Qy 1391 CCGTGCAGCCATCGAGCTGCCCGAGAAGAGAGCTGGACGTGAACGATCCAGAGAGC 1450
Db 2600 CAGTACAACCTATACAGCTGCCAGAAAAGGATAGCTGGACTGTCAATGATATACAGAAGT 2659

Qy 1451 TGGTGGGCAAGTGAACTGGGCGCGCCAGCGCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGT 1510
Db 2660 TAGTGGGAAAAATTAAACTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAGGCAATTGT 2719

Qy 1511 GCAAGCTGCTCGCGGCGCCAAGGCCCTGACCGTGCCCCTGACCGGAGGAGGCCG 1570
Db 2720 GTAAACTCCTTAGGGGAACCAAAGCCATTAACAGACATAGTACCACTAACTGAAGAGCAG 2779

Qy 1571 AGCTGGGAGCTGGCCGAGAACCCGCGAGAACCTCCGCCGGAGCCCCGTGACGCGGTCTACTACG 1630
Db 2780 AATTAGAATTGGCAGAAAACAGGGAAATTTTAAAAGAACCAGTACATGGAGTATATTATG 2839

Qy 1631 ACCCCAGCAAGGAACCTGTGGCCGCGAGATCGAAGCAGGGCCACGGACCAGGGCGACCTACC 1690
Db 2840 ACCCATCAAAAGACTTAATAGCTGAAATACAGAAACAGGGGCATGGCCAATGGACATATC 2899

Qy 1691 AGATCTACCAGGAGACCCCTTCAAGAACCTGAAGACCTGGAGACTACCGCCAAGATGCGGCACCG 1750

Qy 1751 CCCACCAACCAACGTGAAGCAGCAGCTGACCGAGCCGTGCAGAAGATCGCCATGGAGAGCA 1810
Db 2900 AANTTACCAAGAACCATTCAAAAATCTGAAAAATCAGGAAAGTATGCAAAAAAGGAGGACTG 2959
Db 2960 CCCACCTAATGACGTAAAACAGTTAACAGAGGCAGTGCAAAAGATAGCCATGGAAAGCAA 3019

Qy 1811 TCGTGATCTGGGGCAAGAACACCCCCAAGTTCCGCTGCCCCATCCAGAAGAGAGACCTGGGAGA 1870
Db 3020 TAGTAATATGGGAAAAGACTCCTAAATTTAGATTACCCATCCAGAAAGAAACATGGGAAA 3079

Qy 1871 CCTGGTGGACCGACTACTGCAGGCCACCTGGATCCCCGAGTGGGGAGTTCGTGAACACCC 1930
Db 3080 CATGGTGGACAGACTATTGGCAAGCCACCTGGATTCCTGAGTGGGGAGTTTGTTAATACCC 3139

Qy 1931 CCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGGAGGAGCCCATCATCCGCGCCGAGACCT 1990
Db 3140 CTCCCCCAGTAAAATTATGTACCAGCTAGAGAAAAGAACCCATATTAGGAGCAGAGACTT 3199

Qy 1991 TCTACGTGCAGCGCGCCGCAACGTGAGCGTGGAGCTGCAGGCCGGCTACGTGACCG 2050
Db 3200 TCTATGTGCAGCGATGGAGCAGTAATCGGGAAAACTAAAATAGGAAAAAGCAGGTATGTTACTG 3259

Qy 2051 ACCGGGGCCCGCGAGAAGATCGTGAGCGTGACGAGCTGGCACCACCACCAGAAGACCGAGCTGC 2110
Db 3260 ACAGAGGAAGGCAGCAGAAAAATTGTTTCTCTAACTGAAACACAATCAGAGACTGAATTAC 3319

Qy 2111 AGGCCATCCAGCTGCCCTGCAGGACAGCAGCAGCAGCAGCGCGAGCATGCTGACCGAGCC 2170
Db 3320 AAGCAATTCAGCTAGCTTTACAGGATTCAGGATCAGAAGTAAACATAGTAAACAGACTCAC 3379

Qy 2171 AGTACGCCCTGGGCATCATCCAGCCCCAACAAGAGCGAGCTGGTGAACC 2230
Db 3380 AGTATGCATTAGGAATCATTCAAGCACAACCAGATAGCAGCCAGTACAAT 3439

Qy 2231 AGATCATCGAGCAGCTGATCAAGAAGGAAGTGTACCTGAGCTGGGTGCCCGCCACA 2290
Db 3440 AAATAATAGAACAGTTAATACAAAAAGGAAAAGAGTCTACCTGTCATGGGTACCACACATA 3499

Qy 2291 AGGGCGGCCAACGACAGATCGACAAGCTGGTGAGCAAGCTCCCGCAAGGTGC 2350
Db 3500 AAGGAATTGGAGGAAATGAACAGTGAATAATTTAGTAAGTAGTGGAATCAGGAGAGTGC 3559

Qy 2351 TGTTCCTGACGGCATCGAT 2370
Db 3560 TGTTTCTAGATGGAATAGAT 3579

Search completed: April 10, 2004, 16:22:37
Job time : 129.622 secs

```
          GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        April 10, 2004, 12:49:38  ; Search time 594.779 Seconds
                                            (without alignments)
                                            15572.780 Million cell updates/sec

Title:         US-09-610-313-30
Perfect score: 2469
Sequence:      1 gtcgacgccaccatggccga.........gggctagcacggtgaattc 2469

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      2475585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters:      4951170

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  Published Applications NA:*
    1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
    2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
    3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
    4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
    5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
    6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
    7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
    8: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
    9: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
   10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
   11: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
   12: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
   13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
   14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
   15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
   16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
   17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
   18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Score	Length	DB	ID	Description
1	2457	99.5		2457	14	US-10-190-435-45	Sequence 45, Appl
2	2457	99.5		2457	14	US-10-190-305A-39	Sequence 39, Appl
3	2434.6	98.6		2457	14	US-10-190-435-44	Sequence 44, Appl
4	2434.6	98.6		2457	14	US-10-190-305A-38	Sequence 38, Appl
5	2401.8	97.3		2445	14	US-10-190-435-43	Sequence 43, Appl
6	2401.8	97.3		2445	14	US-10-190-305A-37	Sequence 37, Appl
7	2394.8	97.0		3930	14	US-10-190-435-9	Sequence 9, Appli
8	2393.2	96.9		3930	14	US-10-190-435-10	Sequence 10, Appl
9	2393.2	96.9		3930	14	US-10-190-435-11	Sequence 11, Appl
10	2393.2	96.9		5184	14	US-10-190-435-58	Sequence 58, Appl
11	2391.2	96.9		5184	14	US-10-190-305A-82	Sequence 82, Appl
12	2362.8	95.7		3531	14	US-10-190-435-13	Sequence 13, Appl
13	2361.2	95.6		3537	14	US-10-190-435-14	Sequence 14, Appl
14	2361.2	95.6		3537	14	US-10-190-435-15	Sequence 15, Appl
15	2360.2	95.6		5145	14	US-10-190-435-12	Sequence 12, Appl
16	2360.2	95.6		5145	14	US-10-190-305A-12	Sequence 12, Appl
17	2349.4	95.2		3624	14	US-10-190-435-47	Sequence 47, Appl
18	2349.4	95.2		3624	14	US-10-190-305A-41	Sequence 41, Appl
19	2301.8	93.2		3607	14	US-10-190-435-48	Sequence 48, Appl
20	2301.8	93.2		3607	14	US-10-190-305A-42	Sequence 42, Appl
21	2283.6	92.5		3597	14	US-10-190-435-46	Sequence 46, Appl
22	2283.6	92.5		3597	14	US-10-190-305A-40	Sequence 40, Appl
23	2149.6	87.1		2472	14	US-10-241-009-32	Sequence 32, Appl
24	2149.6	87.1		2472	14	US-10-190-434B-32	Sequence 32, Appl
25	2149.6	87.1		2472	14	US-10-190-305A-34	Sequence 34, Appl
26	2121.2	85.9		2466	14	US-10-241-009-31	Sequence 31, Appl
27	2121.2	85.9		2466	14	US-10-190-434B-31	Sequence 31, Appl
28	2121.2	85.9		2466	14	US-10-190-305A-33	Sequence 33, Appl
29	2094.4	84.8		2460	14	US-10-241-009-30	Sequence 30, Appl
30	2094.4	84.8		2460	14	US-10-190-434B-30	Sequence 30, Appl
31	2094.4	84.8		2460	14	US-10-190-305A-32	Sequence 32, Appl
32	2093.4	84.8		3564	14	US-10-241-009-13	Sequence 13, Appl
33	2093.4	84.8		3564	14	US-10-241-009-14	Sequence 14, Appl
34	2093.4	84.8		3564	14	US-10-190-434B-13	Sequence 13, Appl
35	2093.4	84.8		3564	14	US-10-190-434B-14	Sequence 14, Appl
36	2092.8	84.8		4716	14	US-10-190-435-17	Sequence 17, Appl
37	2092.8	84.8		4716	14	US-10-190-305A-13	Sequence 13, Appl
38	2089	84.6		3999	14	US-10-241-009-9	Sequence 9, Appli
39	2089	84.6		3999	14	US-10-190-434B-9	Sequence 9, Appli
40	2087.4	84.5		3999	14	US-10-241-009-10	Sequence 10, Appl
41	2087.4	84.5		3999	14	US-10-241-009-11	Sequence 11, Appl
42	2087.4	84.5		3999	14	US-10-190-434B-10	Sequence 10, Appl
43	2087.4	84.5		3999	14	US-10-190-434B-11	Sequence 11, Appl
44	2087.4	84.5		5283	14	US-10-241-009-54	Sequence 54, Appl
45	2087.4	84.5		5283	14	US-10-190-434B-54	Sequence 54, Appl

ALIGNMENTS

```
RESULT 1
US-10-190-435-45
; Sequence 45, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;   APPLICANT: ZUR MEGEDE, Jan
;   APPLICANT: BARNETT, Susan W.
;   APPLICANT: LIAN, Ying
;   APPLICANT: ENGELBRECHT, Susan
;   APPLICANT: VAN RENSBURG, Estrelita J.
;   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;   TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;   FILE REFERENCE: PP18133.003 / 2302-18133
;   CURRENT APPLICATION NUMBER: US/10/190,435
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 319
;   SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45
;   LENGTH: 2457
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: p2Polopt_C
US-10-190-435-45

Query Match             99.5%;  Score 2457;  DB 14;  Length 2457;
Best Local Similarity   99.5%;
Matches 2457;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 GCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGGCGCCAACATCCTGATGCAGCGCAGC 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGGCGCCAACATCCTGATGCAGCGCAGC 60

Qy     67 AACTTCAAGGGCCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC 120

Qy    127 GCCCGCAACTGCCGCGCCCCCGCCCCCCGGAAGAAGGGCTGCTGGAAGTGCGGCAAGAGGAGGGCCAC 186
```

```
Db  121  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  180
         GCCCGCAACTGCCGCGCCCCGCCAAGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC
Qy  187  CAGATGAAGGACTGCACCGAGCCGCAGCCCAACTTCTTCCGCGAGCCGGCCTTCCCC  246
         ||||||||||||||  |||||||||||||||||||| ||||||||||||
Db  181  CAGATGAAGGACTGCACCGAGCCGCAGCCCAACTTCTTCCGCGAGCCGGCCTTCCCC  240

Qy  247  CAGGGCAAGGCCGGCGAGTTCCCCAGCAGCAGAACCGGCGCCAACAGCCGC      306
         ||||||||||||||||| |||||||||
Db  241  CAGGGCAAGGCCGGCGAGTTCCCCAGCGAGCAGAACCGGCGCCAACAGCCGC      300

Qy  307  GAGCTGCAGGTGCGCGGCGACAACCCCGCAGCGAGCGCCGGCGCCGCCAGGGCCACC  366
         |||||||||||||||||||||||| |||||||||||||||||||
Db  301  GAGCTGCAGGTGCGCGGCGACAACCCCGCAGCGAGCGCCGGCGCCGCCAGGGCCACC  360

Qy  367  CTGAACTTCCCCCAGATCACCGTGTGCAGCGCCGCCCCTGGTCAGCGCCCGTGGGCGGC  426
         |||||||||||||||||||||||||
Db  361  CTGAACTTCCCCCAGATCACCGTGTGCAGCGCCGCCCCTGGTGAGCATCAAGGTGGGCGGC  420

Qy  427  CAGATCAAGGAGGGCCCTGCTGACACCGGCGCCGACACCGGCGCCGACACCGGCCGTGAGATGAGC  486
         ||||||||||||||||||||||||||
Db  421  CTGAACTTCCCCCAGATCACCGTGTGCAGCGCCGCCCCTGGTGAGCATCAAGGTGGGCGGC  480

Qy  487  CTGCCCGGCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC  546
         |||||||||||||||||||||||||
Db  481  CTGCCCGGCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC  540

Qy  547  CAGTACGACCAGATCCTGATGAGATCTGAGCTGCGCCAAGAAGGCCATCGGACCTGCTGTGATC  606
         |||||||||||||||||||||||||||
Db  541  CAGTACGACCAGATCCTGATGAGATCTGAGCTGCGCCAAGAAGGCCATCGGACCTGCTGTGATC  600

Qy  607  GGCCCACCCCGTGAACATCATCGGCCGCCAGTGGCTGCACCCTG      666
         |||||||||||||||||||||||||||
Db  601  GGCCCACCCCGTGAACATCATCGGCCGCCAGTGGCTGCACCCTG      660

Qy  667  AACTTCCCCATCAGCCCGTGAAGCTGCCCGTCGAAGCTGAAGCCCGGCCATGACGGC  726
         ||||||||||||||||||||||||
Db  661  AACTTCCCCATCAGCCCGTGAAGCTGCCCGTCGAAGCTGAAGCCCGGCCATGACGGC  720

Qy  727  CCCAAGGTGAAGCAGTGCCCCTGACCGAGAAGATCAAGGCCCTGACCGCCATCTGC  786
         |||||||||||||||||||||||||
Db  721  CCCAAGGTGAAGCAGTGCCCCTGACCGAGAAGATCAAGGCCCTGACCGCCATCTGC  780

Qy  787  GAGGAGATGAGAAGGAGGCCAAGATCACCAGAGATCGGCCCCGAGAACCCTACAACACC  846
         |||||||||||||||||||||||||||
Db  781  GAGGAGATGAGAAGGAGGCCAAGATCACCAGAGATCGGCCCCGAGAACCCTACAACACC  840

Qy  847  CCGTGTTCGGCCATCAAGAAGAGACAGCACCAGTGGGCGGCAAGCTGGTGGACTTCCGC  906
         |||||||||||||||||||||||||
Db  841  CCGTGTTCGGCCATCAAGAAGAGACAGCACCAGTGGGCGGCAAGCTGGTGGACTTCCGC  900

Qy  907  GAGCTGAACAAGCGCACCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCCGCC  966
         |||||||||||||||||||||||||
Db  901  GAGCTGAACAAGCGCACCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCCGCC  960

Qy  967  GGCCTGAAGAAGAAGAGAGAGTTCGGCAAGTCACCGGCCTTCGGGCGACGCGTG  1026
         |||||||||||||||||||||||||
Db  961  GGCCTGAAGAAGAAGAGAGAGTTCGGCAAGTCACCGGCCTTCGGGCGACGCGTG  1020

Qy  1027  CCCCTGACGAGGAACTTCGGCGAGTACACCGCCATCCCCAGCAGCATCAACAACGAG  1086
          ||||||||||||||||||||||||
Db  1021  CCCCTGACGAGGAACTTCGGCGAGTACACCGCCATCCCCAGCAGCATCAACAACGAG  1080

Qy  1087  ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGGCTGAAGGGCCAGCCCAGC  1146
          |||||||||||||||||||||||||
Db  1081  ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGGCTGAAGGGCCAGCCCAGC  1140

Qy  1147  ATCTTCCAGAGCAGCATGACCGACCAAGATCCTGGAGCGCCTTCCCGCCCGCACCCCGAGATC  1206
          |||||||||||||||||||||||||
Db  1141  ATCTTCCAGAGCAGCATGACCGACCAAGATCCTGGAGCGCCTTCCCGCCCGCACCCCGAGATC  1200

Qy  1207  GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCCCAC  1266
          |||||||||||
```

```
Db  1201  GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCCCAC  1260

Qy  1267  CGGGCCAAGATGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGAC  1326
          ||||||||||||||||||||||||
Db  1261  CGGGCCAAGATGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGAC  1320

Db  1327  AAGAAGCACCAGAAGAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG  1386
          |||||||||||||||||||||||||
Qy  1321  AAGAAGCACCAGAAGAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG  1380

Qy  1387  TGGACCGTGCAGCCCATCGAGCTGCCCGAGAGAGAGCTGACCGTGAACCGTGAACGACATCCAG  1446
          ||||||||||||||||||||||||
Db  1381  TGGACCGTGCAGCCCATCGAGCTGCCCGAGAGAGAGCTGACCGTGAACCGTGAACGACATCCAG  1440

Qy  1447  AAGCTGGTGGGCAAGCTGAACTGGGCGACGAAGGCCAGGATCTACCCCGGCATCAAGGTGCGCCAG  1506
          ||||||||||||||||||||||||
Db  1441  AAGCTGGTGGGCAAGCTGAACTGGGCGACGAAGGCCAGGATCTACCCCGGCATCAAGGTGCGCCAG  1500

Qy  1507  CTGTGCAAGCTGCTCGGCGGGCGCCCAAGCCCCTGACCGACATCGTGCCCCTGACCGAGGAG  1566
          ||||||||||||||||||||||||
Db  1501  CTGTGCAAGCTGCTCGGCGGGCGCCCAAGCCCCTGACCGACATCGTGCCCCTGACCGAGGAG  1560

Qy  1567  GCCGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGAGCCGGTGTCACGCGCGTGTAC  1626
          ||||||||||||||||||||||||
Db  1561  GCCGAGCTGGAGCTGGCCGAGAACCGGCGAGATCCTGCGCGAGCCGGTGTCACGCGCGTGTAC  1620

Qy  1627  TACGACCCCAGCAAGGACCTGGTTGGCCAAGATCCAGAAGCAGCTCAGACCAGTGACC  1686
          ||||||||||||||||||||||||
Db  1621  TACGACCCCAGCAAGGACCTGGTTGGCCAAGATCCAGAAGCAGCTCAGACCAGTGACC  1680

Qy  1687  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAAGATACGGCCAAGATGCGC  1746
          ||||||||||||||||||||||||
Db  1681  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAAGATACGGCCAAGATGCGC  1740

Qy  1747  ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGCGCCGTGCAGAAGATCGCCATGGAG  1806
          ||||||||||||||||||||||||
Db  1741  ACCGCCCACACCAACGACGTGAAGCAGCTGACCGAGCGCCGTGCAGAAGATCGCCATGGAG  1800

Qy  1807  AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCGTTCCGCCTGCTGCCCATCCAGAAGGAGACCTGG  1866
          ||||||||||||||||||||||||
Db  1801  AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCGTTCCGCCTGCTGCCCATCCAGAAGGAGACCTGG  1860

Qy  1867  GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCAGTGGGAGTTCGTGAAC  1926
          ||||||||||||||||||||||||
Db  1861  GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCAGTGGGAGTTCGTGAAC  1920

Qy  1927  ACCCCCCCTGTTGAAGCTGTGTGTACCAGCTGGTGTACCAGCTGGAAGGAGCAGGGCCCAG  1986
          ||||||||||||||||||||||||
Db  1921  ACCCCCCCTGTTGAAGCTGTGTGTACCAGCTGGTGTACCAGCTGGAAGGAGCAGGGCCCAG  1980

Qy  1987  ACCTTCTACGTGACGGGCCCGCAGAAGATCGGCAAGGCCCGCTACTGG  2046
          ||||||||||||||||||||||||
Db  1981  ACCTTCTACGTGACGGGCCCGCAGAAGATCGGCAAGGCCCGCTACTGG  2040

Qy  2047  ACCGACCGGGGCCGGCCGCAGAAGATCGTGAGCTGAGCCTGACCGAGAGACCGAG  2106
          ||||||||||||||||||||||||
Db  2041  ACCGACCGGGGCCGGCCGCAGAAGATCGTGAGCTGAGCCTGACCGAGAGACCGAG  2100

Qy  2107  CTGGAGGCCATCAGCTGCCCTGCCCTGCAGAGCCAGCGAGGTGAACATCGTGACCGAC  2166
          ||||||||||||||||||||||||
Db  2101  CTGGAGGCCATCAGCTGCCCTGCCCTGCAGAGCCAGCGAGGTGAACATCGTGACCGAC  2160

Qy  2167  AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTG  2226
          ||||||||||||||||||||||||
Db  2161  AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTG  2220

Qy  2227  AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC  2286
          ||||||||||||||||||||||||
Db  2221  AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC  2280

Qy  2287  CACAAGGGCATCGGCGGCAACGAGATCGGCGGCATCGGTGCAAGGGCATCCGCAAG  2346
          ||||||||||||||||||||||||
Db  2281  CACAAGGGCATCGGCGGCAACGAGATCGGCGGCATCGGTGCAAGGGCATCCGCAAG  2340
```

```
RESULT 2
US-10-190-305A-39
; Sequence 39, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt_C
US-10-190-305A-39

Query Match        99.5%;  Score 2457;  DB 14;  Length 2457;
Best Local Similarity  100.0%;  Pred. No. 0;
Matches  2457;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      7 GCCACCATGGCCGAGGCCATGAGCGCCAGCGCCAGCGCCAACATCCTGATGCAGCGCAGC 66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCCACCATGGCCGAGGCCATGAGCGCCAGCGCCAGCGCCAACATCCTGATGCAGCGCAGC 60

Qy     67 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 120

Qy    127 GCCCGCAACTGCGCCCCGCGCCCCCGCAAGAGAGGCTGCTGGAAGTGCGGCAAGAGGGCAC 186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCCCGCAACTGCGCCCCGCGCCCCCGCAAGAGAGGCTGCTGGAAGTGCGGCAAGAGGGCAC 180

Qy    187 CAGATGAAGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCC 246
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CAGATGAAGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCC 240

Qy    247 CAGGGCAAGGCCCGAGTTCCCCAGCGAGCAGCAGAACCGGCCAACAGCCCCACCAGCCGC 306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGGGCAAGGCCCGAGTTCCCCAGCGAGCAGCAGAACCGGCCAACAGCCCCACCAGCCGC 300

Qy    307 GAGCTGCAGGTGCGCGGACAACCCCCGCAGCGCCGGCGGCCGGCCGCCGGGTGGGGCGGC 366
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GAGCTGCAGGTGCGCGGACAACCCCCGCAGCGCCGGCGGCCGGCCGCCGGGTGGGGCGGC 360

Qy    367 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGGGGC 426
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGGGGC 420

Qy    427 CAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTCCTGGAGGAGATGAGC 486
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTCCTGGAGGAGATGAGC 480

Qy    487 CTGCCCGGCCGCTGGAAGCCCAAGATGATCGGCGGGCATCGGCGGCTTCATCAAGGTGCGC 546
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGCCCGGCCGCTGGAAGCCCAAGATGATCGGCGGGCATCGGCGGCTTCATCAAGGTGTGCGC 540

Qy    547 CAGTACGACCAGATCCTGATCGATGAGATCTGGCGGCAAGAGAAGGCCATCGGCACCGTGCTGCTGATC 606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGTACGACCAGATCCTGATCGATGAGATCTGGCGGCAAGAGAAGGCCATCGGCACCGTGCTGCTGATC 600

Qy    607 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTCGCACCCTG 666
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGCCCCACCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTCGCACCCTG 660

Qy    667 AACTTCCCCATCAGCCCATGAGACGACCGTGCCCGTGAAGCTGAAGCCTGAAGTGAAGCATGGCGACGGC 726
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AACTTCCCCATCAGCCCATGAGACGACCGTGCCCGTGAAGCTGAAGCCTGAAGTGAAGCATGGCGACGGC 720

Qy    727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGACATCTGC 786
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGACATCTGC 780

Qy    787 GAGGAGATGGAGAAGGAGGGCAAGATCAGCAGCACCAAGAACCCCTACAACACC 846
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGATGGAGAAGGAGGGCAAGATCAGCAGCACCAAGAACCCCTACAACACC 840

Qy    847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 906
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC 900

Qy    907 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC 966
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC 960

Qy    967 GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG 1026
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG 1020

Qy   1027 CCCCTGGACGAGGACTTCCGCAAGTACACACCGCCTTCACCATCCCCAGCATCAACAACGAG 1086
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCCTGGACGAGGACTTCCGCAAGTACACACCGCCTTCACCATCCCCAGCATCAACAACGAG 1080

Qy   1087 ACCCCGCGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 ACCCCGCGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1140

Qy   1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1206
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1200

Qy   1207 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1266
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1260

Qy   1267 CGCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1326
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CGCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGAC 1320

Qy   1327 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCAGACAAG 1386
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCAGACAAG 1380

Qy   1387 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG 1446
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG 1440

Qy   1447 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAG 1506
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAG 1500

Qy   1507 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1566
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy   1567 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1626
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTAC 1620

Qy   1627 TACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGGCCAGTGGACC 1686

Qy   2347 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGTGATCTCACCAGTACATGGACGACCTG 2406
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGTGATCTCACCAGTACATGGACGACCTG 2400

Qy   2407 TACGTGGGCAGCGGCGGCGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGT 2463
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 TACGTGGGCAGCGGCGGCGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGT 2457
```

; SEQ ID NO 44
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt.YM_C
US-10-190-435-44

Query Match   98.6%;   Score 2434.6;   DB 14;   Length 2457;
Best Local Similarity   99.6%;   Pred. No. 0;
Matches 2453;   Conservative   4;   Indels   6;   Gaps   1;
                Mismatches   0;   Conservative   0;

Qy   1 GTCGACGGCCACCATGGCCGGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG   60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTCGACGGCCACCATGGCCGGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG   60

Qy   61 CGAGCAACTTCAAGGGCCCAAGCGCGATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC   120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CGAGCAACTTCAAGGGCCCAAGCGCGATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC   120

Qy 121 CACATCGCCCGCAACTGCCGCGCCCCCGCAAGAAGGCCGTGCTGGAAGTGCGCAAGGAG   180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CACATCGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCCGTGCTGGAAGTGCGCAAGGAG   180

Qy 181 GGCCACCAGATGAAGGACTGCACCGAGCGCCCCAGGCCAGGAGCGGCGCGGAGCCTGGCC   240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GGCCACCAGATGAAGGACTGCACCGAGCGCCCCAGGCCAGGAGCGGCGCGGAGCCTGGCC   240

Qy 241 TTCCCCAGGGCAAGGCCGGTGCGCGGCGACAACCCCGCCAGCGAGCGCCAGCGCCAG   300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 TTCCCCAGGGCAAGGCCGGTGCGCGGCGACAACCCCGCCAGCGAGCGCCAGCGCCAG   300

Qy 301 AGCCGCGAGCTGCAGGTGCGGCGGCGACAACATCGTGACCGAC   360
         |||||||||||||||||||||||||||||||||||||||||||
Db 301 AGCCGCGAGCTGCAGGTGCGGCGGCGACAACATCGTGACCGAC   360

Qy 361 GGCACCCTGAACTTCCCCGAGATCACCCTGTGCAGCGGCGCCCCC   420
         ||||||||||||||||||||||||||||||||||||||||||||
Db 361 GGCACCCTGAACTTCCCCGAGATCACCCTGTGCAGCGGCGCCCCC   420

Qy 421 GGCGGCCAGATCAAGGAGGACAGCGAGCTGTACCTGGACACCGGCATCGGCGCTTCATCAAG   480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 GGCGGCCAGATCAAGGAGGACAGCGAGCTGTACCTGGACACCGGCATCGGCGCTTCATCAAG   480

Qy 481 ATCTCGGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCATCGGCGGCACCGTG   540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 ATCTCGGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCATCGGCGGCACCGTG   540

Qy 541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG   600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG   600

Qy 601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC   660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGC   660

Qy 661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG   720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG   720

Qy 721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC   780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC   780

Qy 781 ATCTCGGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC   840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 ATCTCGGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC   840

Qy 841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC   900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC   900

Qy 901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC   960

Db 1621 TACGACCCCCAGGAAGGACCTGCTGTGGCCGAGATCCCAGAAGCAGGGCGCACGACCAGCTGGACC   1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 1687 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC   1746
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC   1740

Qy 1747 ACCGGCCCACCAACCACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG   1806
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 ACCGGCCCACCAACCACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG   1800

Qy 1807 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG   1866
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGG   1860

Qy 1867 GAGACCTGGTGACCGACTACTGCGAGCACCTGGAGAAGGAGCCCATCGGGGCGGCGGAAC   1926
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 GAGACCTGGTGACCGACTACTGCGAGCACCTGGAGAAGGAGCCCATCGGGGCGGCGGAAC   1920

Qy 1927 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCTCGGGCGGCGAG   1986
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCTCGGGCGGCGAG   1980

Qy 1987 ACCTTCTACGTGGACGGCGCGGCCAACAGGGCGCCAAGATCGGCAAGGCCGGCTACGTG   2046
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 ACCTTCTACGTGGACGGCGCGGCCAACAGGGCGCCAAGATCGGCAAGGCCGGCTACGTG   2040

Qy 2047 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG   2106
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 ACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG   2100

Qy 2107 CTGCAGGCCATCCAGTTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC   2166
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 CTGCAGGCCATCCAGTTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGAC   2160

Qy 2167 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTG   2226
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 AGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGCTGGTG   2220

Qy 2227 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCCTGGGTGCCCGCC   2286
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGCCTGGGTGCCCGCC   2280

Qy 2287 CACAAGGGCATCGGCGGCAACGAGCAGGTAGATCGACAACTGGTGACCAAGGGCATCGCCAAG   2346
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 CACAAGGGCATCGGCGGCAACGAGCAGGTAGATCGACAACTGGTGACCAAGGGCATCGCCAAG   2340

Qy 2347 GTGCTGTTCCTGGACGGCATCGCGGCATCTGATCTACCAGTACATGGACCACCTG   2406
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 GTGCTGTTCCTGGACGGCATCGCGGCATCTGATCTACCAGTACATGGACCACCTG   2400

Qy 2407 TACGTGGGCAGCGGCGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT   2463
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 TACGTGGGCAGCGGCGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT   2457

RESULT 3
US-10-190-435-44
; Sequence 44, Application US/10190435
; Publication No. US2003014348A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; TITLE OF INVENTION: POLYPEPTIDES,
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0

```
Db    901  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
           TTCCGCGAGCTGAACAAGCGCACCCAGGAGGTGCAGCTGGGCATCCCCCAC    960
Qy    961  CCCGCCGGCCTGAAGCAGAAGAAGAGCGGTGACCGTGCTGGACGTGGGCGAGCGCCTACTTC
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  CCCGCCGGCCTGAAGCAGAAGAAGAGCGGTGACCGTGCTGGACGTGGGCGAGCGCCCTACTTC   1020
Qy   1021  AGCGTGCCCCTGGACGACGAGCGCACTTCCGCAAGCAAGTACACCACCCCCAGCATCAAC   1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  AGCGTGCCCCTGGACGACGAGCGCACTTCCGCAAGCAAGTACACCACCCCCAGCATCAAC   1080
Qy   1081  AACGAGACCCCGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC   1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  AACGAGACCCCGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC   1140
Qy   1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC   1200
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCC   1200
Qy   1201  GAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGGC   1260
           |||||||||||||||||                  |||||||||||||||||||||||||
Db   1201  GAGATCGTGATCTACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGGC   1254
Qy   1261  CAGCACCGCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACC   1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255  CAGCACCGCGCCAAGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACC   1314
Qy   1321  CCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCC   1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315  CCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCC   1374
Qy   1381  GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC   1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1375  GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC   1434
Qy   1441  ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG   1500
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1435  ATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG   1494
Qy   1501  CGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACC   1560
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1495  CGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACC   1554
Qy   1561  GAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGC   1620
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1555  GAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGC   1614
Qy   1621  GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAG   1680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1615  GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAG   1674
Qy   1681  TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAG   1740
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1675  TGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAG   1734
Qy   1741  ATCGCACCGCCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCC   1800
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1735  ATGGCACCGCCACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCC   1794
Qy   1801  ATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG   1860
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1795  ATGGAGAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG   1854
Qy   1861  ACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC   1920
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1855  ACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTC   1914
Qy   1921  GTGAACACACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC   1980
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1915  GTGAACACACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGC   1974
Qy   1981  GCCGAGACCTTCTACGTGGACGGCGCCCGCCAACCGCGAGACCCAAGATCGGCAAGGCCGGC   2040
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1975  GCCGAGACCTTCTACGTGGACGGCGCCCGCCAACCGCGAGACCCAAGATCGGCAAGGCCGGC   2034
Qy   2041  TACGTGACCGACCGGGGCCGGCCGCAGAAGATCGTGAGCGCTGAGCCAGGACCACCAGGAAG   2100
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2035  TACGTGACCGACCGGGGCCGGCCGCAGAAGATCGTGAGCGCTGAGCCTGACCGACCAACCAGAAG   2094
Qy   2101  ACCGAGCTGCAGGCCATCAGCTGGCCCTGCCAGGACACGAGCGGCAGCGGGTGAACATCGTG   2160
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2095  ACCGAGCTGCAGGCCATCAGCTGGCCCTGCCAGGACACGAGCGGCAGCGGGTGAACATCGTG   2154
Qy   2161  ACCGACGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGC   2220
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2155  ACCGACGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAG   2214
Qy   2221  CTGGTGAACCAGACATCATGCAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG   2280
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2215  CTGGTGAACCAGATCATGCAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTG   2274
Qy   2281  CCCGCCACAAGGGCATCGGCGGCAACGACATCGGCAACGAGCAGATCGTGATCTACCAGTACATGGACATC   2340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2275  CCCGCCACAAGGGCATCGGCGGCAACGAGCAGATCGTGATCTACCAGTACATGGACATC   2334
Qy   2341  CGCAAGTGCTGTTCCTGGACGGCATCGATGCGGCATCGATGCGGCATCTACCAGTACATGGAC   2400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2335  CGCAAGTGCTGTTCCTGGACGGCATCGATGCGGCATCGATGCGGCATCTACCAGTACATGGAC   2394
Qy   2401  GACCTTACGTGGGCAGCGGCGGCCCTAGGATCAGATCGATTAAAAGCTTCCCGGGGCTAGCACC   2460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2395  GACCTGTACGTGGGCAGCGGCGGCCCTAGGATCAGATCGATTAAAAGCTTCCCGGGGCTAGCACC   2454
Qy   2461  GGT   2463
           |||
Db   2455  GGT   2457
```

RESULT 4
US-10-190-305A-38
; Sequence 38, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2202-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Polopt.YM_C
US-10-190-305A-38

Query Match   98.6%; Score 2434.6; DB 14; Length 2457;
Best Local Similarity   99.6%; Pred. No. 0;
Matches 2453; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

```
Qy      1  GTCGGACGGCCACCATGGCCGAGGCCATGAGCCCAGCGCCAACATCCTGATGCAG   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GTCGACCGCCACCATGGCCGAGGCCATGAGCCCAGCGCCAACATCCTGATGCAG   60
Qy     61  CGCAGCAACTTCAAGGCCCAAGCCCATCATCAAGTGCTTCAACTGCGGCAAGGGGC   120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CGCAGCAACTTCAAGGCCCAAGCCCATCATCAAGTGCTTCAACTGCGGCAAGGGGC   120
Qy    121  CACATCGCCGGCAACTGCCGCGCCCCCGCCCCCCGCCAAGAAGGGGCTGCTGGAAGTGCGGCAAGGAG   180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CACATCGCCGGCAACTGCCGCGCCCCGCCCCCCGCCAAGAAGGGGCTGCTGGAAGTGCGGCAAGGAG   180
```

```
1261 Qy CAGCACCGCGCCAAGATCGAGGAGCTGCTGCGCTGCTGGGCTTCACCACC 1320
        |||||||||||||||||||||||||||||||||||||||||||||||||
1255 Db CAGCACCGCGCCAAGATCGAGGAGCTGCTGCGCTGCTGGGCTTCACCACC 1314

1321 Qy CCCGACAAGAGCACCAGAGGAGGAGCCCCTTCGTGATGGGCTACGAGCTGCACCCC 1380
        |||||||||||||||||||||||||||||||||||||||||||||||||
1315 Db CCCGACAAGAGCACCAGAGGAGGAGCCCCTTCGTGATGGGCTACGAGCTGCACCCC 1374

1381 Qy GACAAGTGGACCGTGCAGTCGCCCAGCGTGCCCCAGAACGAGAGCTGACCGTGAACGAC 1440
        |||||||||||||||||||||||||||||||||||||||||||||||||
1375 Db GACAAGTGGACCGTGCAGTCGCCCAGCGTGCCCCAGAACGAGAGCTGACCGTGAACGAC 1434

1441 Qy ATCCAGAAGCTGGTGGCAGCTGAACTGGGCCAGCCAGATCTCACCCCGCATCAAGGTG 1500
        |||||||||||||||||||||||||||||||||||||||||||||||||
1435 Db ATCCAGAAGCTGGTGGCAGCTGAACTGGGCCAGCCAGATCTCACCCCGCATCAAGGTG 1494

1501 Qy CGCCAGCTGTGCAAGCTGCTGCGGCGGCCCAAGGCCCTGACCGACATCGTGCCCCTGACC 1560
        |||||||||||||||||||||||||||||||||||||||||||||||||
1495 Db CGCCAGCTGTGCAAGCTGCTGCGGCGGCCCAAGGCCCTGACCGACATCGTGCCCCTGACC 1554

1561 Qy GAGGAGGCCGAGCTGGAGCTGCCGAGAACCGGCGAGATCCTGCGCGAGCCGTGCACGGC 1620
        |||||||||||||||||||||||||||||||||||||||||||||||||
1555 Db GAGGAGGCCGAGCTGGAGCTGCCGAGAACCGGCGAGATCCTGCGCGAGCCGTGCACGGC 1614

1621 Qy GTGTACTACGACCCCAGCAAGGACCTGGTGCCGAGATCGAGAGCAGGGCCACGACCAG 1680
        |||||||||||||||||||||||||||||||||||||||||||||||||
1615 Db GTGTACTACGACCCCAGCAAGGACCTGGTGCCGAGATCGAGAGCAGGGCCACGACCAG 1674

1681 Qy TGGACTACCAGATTCACCAGGAGCCCTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1740
        |||||||||||||||||||||||||||||||||||||||||||||||||
1675 Db TGGACTACCAGATTCACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1734

1741 Qy ATGGCACCGCCACCGAGCAGTGAGGCCGTGAGCCGGCCGTGCGAGAAGATCGCC 1800
        |||||||||||||||||||||||||||||||||||||||||||||||||
1735 Db ATGGCACCGCCACCGAGCAGTGAGGCCGTGAGCCGGCCGTGCGAGAAGATCGCC 1794

1801 Qy ATGGAGAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1860
        |||||||||||||||||||||||||||||||||||||||||||||||||
1795 Db ATGGAGAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAG 1854

1861 Qy ACCTGGGAGACCTGGTGGCACTGGACCACTGGCAGGCCACTGGAGTCCCGAGTGGGAGTTC 1920
        |||||||||||||||||||||||||||||||||||||||||||||||||
1855 Db ACCTGGGAGACCTGGTGGCACTGGACCACTGGCAGGCCACTGGAGTCCCGAGTGGGAGTTC 1914

1921 Qy GTGAACACCCCCCCGGTGAAGCTGTGTGGTACCAGCTGGAAGGAGCCCATCATCGGC 1980
        |||||||||||||||||||||||||||||||||||||||||||||||||
1915 Db GTGAACACCCCCCCGGTGAAGCTGTGTGGTACCAGCTGGAAGGAGCCCATCATCGGC 1974

1981 Qy GCCGAGACCTTCTACGTGACGAGCCGGCGCCGCGAGACCAAGACCGGAGATCGGCAAGGCCGGC 2040
        |||||||||||||||||||||||||||||||||||||||||||||||||
1975 Db GCCGAGACCTTCTACGTGACGAGCCGGCGCCGCGAGATCGGCAAGATCGGCAAGGCCGGC 2034

2041 Qy TACGTGACCGACCGGGGCCATCCAGCTGACGTGAGCGTGAGCGTGACCGCACCAACCAGAAG 2100
        |||||||||||||||||||||||||||||||||||||||||||||||||
2035 Db TACGTGACCGACCGGGGCCATCCAGCTGACGTGAGCGTGAGCGTGACCGCACCAACCAGAAG 2094

2101 Qy ACCGAGCTGAGCGTGCGCATCCAGCTGGCCAGTGAAACATCGTG 2160
        |||||||||||||||||||||||||||||||||||||||||||||||||
2095 Db ACCGAGCTGAGCGTGCGCATCCAGCTGGCCAGTGAAACATCGTG 2154

2161 Qy ACCGACAGCCAGTACGCCCTGGCCATCATCGGGGTGGAAGGGCGAG 2220
        |||||||||||||||||||||||||||||||||||||||||||||||||
2155 Db ACCGACAGCCAGTACGCCCTGGCCATCATCGGGGTGGAAGGGCGAG 2214

2221 Qy CTGGTGAACCAGATCATCGAGCGACTGGATCAAGAAGAGGTGTACCTGAGCTGGGTG 2280
        |||||||||||||||||||||||||||||||||||||||||||||||||
2215 Db CTGGTGAACCAGATCATCGAGCGACTGGATCAAGAAGAGGTGTACCTGAGCTGGGTG 2274

2281 Qy CCCGCCCAAGGGCATCGGCGACGAGCAACGAGACCGAGATCGTGTGAGCCAAGGGCATC 2340
        |||||||||||||||||||||||||||||||||||||||||||||||||
2275 Db CCCGCCCAAGGGCATCGGCGACGAGCAACGAGATCGACAACGTGTGAGCAAGGGCATC 2334

2341 Qy CGCAAGGTGCTGTTCCTGGACGGCATCGATGCGGCATGCGTGATCTGATCTACCACCAGTACATGGAC 2400
```

```
 181 Qy GGCCACCGAGATGAAGGACTGCACCGAGGCGCCAACTTCTTCGCGAGGACCTGGCC 240
        ||||||||||||||||||||||||||||||||||||||||||||||
 181 Db GGCCACCAGAGATGAAGGACTGCACCGAGGCGCCAACTTCTTCGCGAGGACCTGGCC 240

 241 Qy TTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGCAGCAGAACCGCGCCAACAGCCCCACC 300
        ||||||||||||||||||||||||||||||||||||||||||||||
 241 Db TTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGCAGCAGAACCGCGCCAACAGCCCCACC 300

 301 Qy AGCCGCGAGCTGCAGGTGCGCGGCGACAACCCCGCAGCGAGCGAGCGCCGAGCGCCAG 360
        ||||||||||||||||||||||||||||||||||||||||||||||
 301 Db AGCCGCGAGCTGCAGGTGCGCGGCGACAACCCCGCAGCGAGCGAGCGCCGAGCGCCAG 360

 361 Qy GGCACCCTGAACTTCCCCCAGATCACCCTGTGCGCAGGCCCCTGGTGCAGCATCAAGGTG 420
        ||||||||||||||||||||||||||||||||||||||||||||||
 361 Db GGCACCCTGAACTTCCCCCAGATCACCCTGTGCGCAGGCCCCTGGTGAGCATCAAGGTG 420

 421 Qy GGCGGCCAGATCAAGGAGGCCCTGCTGACACCGGCGCCGACACCGTGCTGACCGTGGAGGAG 480
        ||||||||||||||||||||||||||||||||||||||||||||||
 421 Db GGCGGCCAGATCAAGGAGGCCCTGCTGACACCGGCGCCGACACCGTGCTGACCGTGGAGGAG 480

 481 Qy ATGAGCCTGCCCCGGCAAGTGGAAGCCCAAGATGATGGAAGCCCAAGATGATCGGCTTTCATCAAG 540
        ||||||||||||||||||||||||||||||||||||||||||||||
 481 Db ATGAGCCTGCCCCGGCAAGTGGAAGCCCAAGATGATGGAAGCCCAAGATGATCGGCTTTCATCAAG 540

 541 Qy GTGCGCCAGTACGACACCGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTG 600
        ||||||||||||||||||||||||||||||||||||||||||||||
 541 Db GTGCGCCAGTACGACACCGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTG 600

 601 Qy CTGATCGGCCCCACCCCCGTGAACATCGGCCGCAACATGCTGACCCAGTCGGCTGC 660
        ||||||||||||||||||||||||||||||||||||||||||||||
 601 Db CTGATCGGCCCCACCCCCGTGAACATCGGCCGCAACATGCTGACCCAGTCGGCTGC 660

 661 Qy ACCCTGAACTTCCCCATCAGCCCCATCGAGACAGCCGTGAAGCTGAAGCCGGGATG 720
        ||||||||||||||||||||||||||||||||||||||||||||||
 661 Db ACCCTGAACTTCCCCATCAGCCCCATCGAGACAGCCGTGAAGCTGAAGCCGGGCATG 720

 721 Qy GACGGCCCAAGGTGAAGCAGTGACGGCCCTGACCGAGAAGATCAAGGCCCCGAGAACCCCTAC 780
        ||||||||||||||||||||||||||||||||||||||||||||||
 721 Db GACGGCCCAAGGTGAAGCAGTGACGGCCCTGACCGAGAAGATCAAGGCCCCGAGAACCCCTAC 780

 781 Qy ATCTGCGAGCAGTGGCAGGAGGCCAAGATCACCAACCAAGATCGGCCCCGAGAACCCCTAC 840
        ||||||||||||||||||||||||||||||||||||||||||||||
 781 Db ATCTGCGAGCAGTGGCAGGAGGCCAAGATCACCAACCAAGATCGGCCCCGAGAACCCCTAC 840

 841 Qy AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGGAC 900
        ||||||||||||||||||||||||||||||||||||||||||||||
 841 Db AACACCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGGAC 900

 901 Qy TTCCGCGAGCTGAACAAGGCCGCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
        ||||||||||||||||||||||||||||||||||||||||||||||
 901 Db TTCCGCGAGCTGAACAAGGCCGCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960

 961 Qy CCCGCCGGCCTGAAGGAGTGCGCCGTGCTGACCGTGCTGGGCGGACGCCGCTACTTC 1020
        ||||||||||||||||||||||||||||||||||||||||||||||
 961 Db CCCGCCGGCCTGAAGGAGTGCGCCGTGCTGACCGTGCTGGGCGGACGCCGCTACTTC 1020

1021 Qy AGCGTGCCCCTGGACGTTCCGCGATACCAGTACACCCCCAGCATCAAAC 1080
        ||||||||||||||||||||||||||||||||||||||||||||||
1021 Db AGCGTGCCCCTGGACGTTCCGCGATACCAGTACACCCCCAGCATCAAAC 1080

1081 Qy AACGAGACCCCCGCCGCATCCGCTACCAGTGCAAGGTACACGTGGGGCGAGGGGCAGC 1140
        ||||||||||||||||||||||||||||||||||||||||||||||
1081 Db AACGAGACCCCCGCCGCATCCGCTACCAGTGCAAGGTACACGTGGGGCGAGGGGCAGC 1140

1141 Qy CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGCGCCCAACCCC 1200
        ||||||||||||||||||||||||||||||||||||||||||||||
1141 Db CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGCGCCCAACCCC 1200

1201 Qy GAGATCGTGATCTACCAGTTACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGGAGATCGGC 1260
        |||||||||||||||||||||||||     ||||||||||||||||||||||||||
1201 Db GAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGGCAGCGACCTGGGAGATCGGC 1254
```

```
Qy   487  CTGCCGGCAAGTGGAAGCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC  546
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db   481  CTGCCGGCAAGTGAAGCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGC   540

Qy   547  CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATGGCCACCTGTGCTGATC  606
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db   541  CAGTACGTGAAGCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATGGCCACCTGTGCTGATC  600

Qy   607  GGCCCCACCCCCGTGAACATCATCGGCCGCCAACATGTCGCCAGCTGGGCTGCACCCTG  666
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601  GGCCCCACCCCCGTGAACATCATCGGCCGCCAACATGTGCTGACCCAGCTGGGCTGCACCCTG  660

Qy   667  AACTTCCCATCAGCCCATGCGCCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC  726
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661  AACTTCCCATCAGCCCATGCGCCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGC  720

Qy   727  CCCAAGGTGAAGCAGTGGCCCTGACCGGAGAGATCAAGGCCCTGACCGCCATCTGC  786
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721  CCCAAGGTGAAGCAGTGGCCCTGACCGGAGAGATCAAGGCCCTGACCGCCATCTGC  780

Qy   787  GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACC  846
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781  GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACC  840

Qy   847  CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGTGGACTTCCGC  906
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841  CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGTGGACTTCCGC  900

Qy   907  GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC  966
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901  GAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCC  960

Qy   967  GGCCTGAAGAAGAAGAAGAGCGTGCTGACCGTGCTGGACGTGGGCGACGCCCTACTTCAGCGTG  1026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961  GGCCTGAAGAAGAAGAAGAGCGTGCTGACCGTGCTGGACGTGGGCGACGCCCTACTTCAGCGTG  1020

Qy  1027  CCCCTGGACGAGGACTTCCGCCAGTACCAGTACAACGTGCTGCCCCAGGCTGAAGGGCAGCGCCAGC  1086
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  CCCCTGGACGAGGACTTCCGCCAGTACCAGTACAACGTGCTGCCCCAGGCTGAAGGGCAGCGCCAGC  1080

Qy  1087  ACCCCCGGCATCCAGCCCAGCCAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCGAGATC  1146
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  ACCCCCGGCATCCAGCCCAGCCAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCGAGATC  1140

Qy  1147  ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCAAGGAGGAGGCCACATC  1206
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCAAGGAGGAGGCCACATC  1200

Qy  1207  GTGATCTACCAGTACATGATGGACGACCTGTACGTGGGCAAGAGTGCGCAAGGAGGGCCAGCAC  1266
          ||||||||||||||||||    | |||||||||||||||||||||||||||||||||||||
Db  1201  GTGATCTACCA------GGCCCCCTGTACGTGGGCAAGAGTGCGCAAGGAGGGCCAGCAC  1254

Qy  1267  CGGCGGCCAAGATCGAGGAGTTGCGCAAGCACCTGCTGCGCGGCTTCACCACCCCGAC  1326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1255  CGGCGGCCAAGATCGAGGAGTTGCGCAAGCACCTGCTGCGCGGCTTCACCACCCCGAC  1314

Qy  1327  AAGAAGCACCAGAAGAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGACCAAG  1386
          ||||||||||||||||||||||||||||||||||||            |||||||||
Db  1315  AAGAAGCACCAGAAGAGGAGCCCCCTTCCTGTGCCCAT------CGAGCTGACCAAG  1368

Qy  1387  TGGACCGTGCAGCCCATGCAGCTGGAAGGAGAGCTGGACCGTGAACGACATCCAG  1446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1369  TGGACCGTGCAGCCCATGCAGCTGGAAGGAGAGCTGGACCGTGAACGACATCCAG  1428

Qy  1447  AAGCTGGTGGCAAGCTGACTGGGCCAGATCTACCCCGGCATCAAGGTGCGCCAG  1506
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1429  AAGCTGGTGGCAAGCTGACTGGGCCAGATCTACCCCGGCATCAAGGTGCGCCAG  1488

Qy  1507  CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACGACCCTGTGCCCCTGACCGGAGAGC  1566
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1489  CTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACGACCCTGTGCCCCTGACCGGAGAGC  1548

Qy  1567  GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGGAGCATCCTGCGGAGAACCAGGAG  1626
```

RESULT 5
US-10-190-435-43
; Sequence 43, Application US/10190435
; Publication No. US2003014324 8A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THERROF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p2Pol.opt.YMWM_C
US-10-190-435-43

Query Match   97.3%; Score 2401.8; DB 14; Length 2445;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2438; Conservative 0; Mismatches 0; Indels 7; Gaps 2;

```
Qy    7  GCCACCATGGCCGAGGCCATGAGCCAGGCCAGGCCACCAGCGCCAACATCTGATGCAGCGCAGC  66
         |||||||||||||||||||||||||||||||||         |||||||||||||||||||||
Db    1  GCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCACCAGCGCCAACATCCTGATGCAGCGCAGC  60

Qy   67  AACTTCAAGGGCCCCAAAGTGCTCATCATCAAGTGCTTCAACTGCAAGAAGAGGAGGCCACATC  126
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AACTTCAAGGGCCCCAAAGTGCTCATCATCAAGTGCTTCAACTGCAAGAAGAGGAGGCCACATC  120

Qy  127  GCCCGCAACTGCCGCGCCCCCGCGCCCGCCCCAGCGAGCAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAAG  186
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GCCCCGCAACTGCCGCGCCCCCGCGCCCGCCCCAGCGAGCAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCAC  180

Qy  187  CAGATGAAGGACTGCACCGAGCGGCCGCCAGGCCAGGGCGCCAGGCCGCCGGAGCGGCGCCACC  246
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CAGATGAAGGACTGCACCGAGCGGCCGCCAGGCCAGGGCGCCAGGCCGCCGGAGCGGCGCCACC  240

Qy  247  CAGGGCCAAGGCCCCGCGCCCCCAGCGACCAGCGAGCAGAACCGCGCCAACAGCCGCCACCAGCCGC  306
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  CAGGGCCAAGGCCCCGCGCCCCCAGCGACCAGCGAGCAGAACCGCGCCAACAGCCGCCACCAGCCGC  300

Qy  307  GAGCTGCAGGTGCGGCGGCGACAACCCCCGCGAGCGGCCTGCCTTCCCC  366
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  GAGCTGCAGGTGCGGCGGCGACAACCCCCGCGAGCGGCCTGCCTTCCCC  360

Qy  367  CTGAACTTCCCCCAGATCACCCTGTGGGCAGCGCCCCCTGTGGTGGGCGGC  426
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CTGAACTTCCCCCAGATCACCCTGTGGGCAGCGCCCCCTGTGGTGGGCGGC  420

Qy  427  CAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGC  486
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  CAGATCAAGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGC  480
```

```
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 37
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: p2Pol.opt.YMWM_C
US-10-190-305A-37
```

```
Query Match          97.3%;  Score 2401.8;  DB 14;  Length 2445;
Best Local Similarity 99.2%;  Pred. No. 0;
Matches 2438;  Conservative  0;  Mismatches  7;  Indels 12;  Gaps  2;
```

```
Qy    7 GCCACCATGGCCGAGGCCATGAGCCAGCCACCAGGCCAACATCCTGATGCAGCGGCAGC   66
        ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GCCACCATGGCCGAGGCCATGAGCCAGCCACCAGGCCAACATCCTGATGCAGCGGCAGC   60

Qy   67 AACTTCAAGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC  126
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AACTTCAAGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATC  120

Qy  127 GCCCGCAACTGCCGCGGCCCCCCGCAAGAAGGGCCATCATCGCGGCGAG...  186
        |||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GCCCGCAACTGCCGCGGCCCCCCGCAAGAAGGGCCATCATCGCGGCGAG...  180

Qy  187 CAGATGAAGGACTGCACCGAGCGGTTCCCAGGCCAGGCGAACTTCTTCCGCGAGGACCTG  246
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 CAGATGAAGGACTGCACCGAGCGGTTCCCAGGCCAGGCGAACTTCTTCCGCGAGGACCTG  240

Qy  247 CAGGGCAAGGCCCGAGAGTTCCCCAGCGAGCGCCAGCGCCCGAGCGCCGAGGGCCACC  306
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 CAGGGCAAGGCCCGAGAGTTCCCCAGCGAGCGCCAGCGCCCGAGCGCCGAGGGCCACC  300

Qy  307 GAGCTGCAGTGCGCGGCGACAACCCCGGCGAGGCGCCCCTGGTGCAGCGGCGG  366
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 GAGCTGCAGTGCGCGGCGACAACCCCGGCGAGGCGCCCCTGGTGCAGCGGCGG  360

Qy  367 CTGAACTTCCCCAGATCACCCTGTGCAGCGGCCCCCTGGTGCGAGCAGATCAAGGTGGGCGGC  426
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 CTGAACTTCCCCAGATCACCCTGTGCAGCGGCCCCCTGGTGCGAGCAGATCAAGGTGGGCGGC  420

Qy  427 CAGATCAAGGAGGCCCTGCTGGACACCGTGCGGGTGGGGATGAGC  486
        |||||||||||||||||||||||||||||||||||||||||||||
Db  421 CAGATCAAGGAGGCCCTGCTGGACACCGTGCGGGTGGGGATGAGC  480

Qy  487 CTGCCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCCTTCATCAAGGTGCGC  546
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 CTGCCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCCTTCATCAAGGTGCGC  540

Qy  547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCCAAGAAGGCCATCATCGAGAAGC  606
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCCAAGAAGGCCATCATCGAGAAGC  600

Qy  607 GGGCCCACCCCGTGAACATCCCGTGAACATCAGCGAGACCTGCCCGTGAAGCTGAAGC  666
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 GGGCCCACCCCGTGAACATCCCGTGAACATCAGCGAGACCTGCCCGTGAAGCTGAAGC  660

Qy  667 AACTTCCCATCACAGCCCCATCGACCCGCCCATCTGC  726
        |||||||||||||||||||||||||||||||||||||
Db  661 AACTTCCCATCACAGCCCCATCGACCCGCCCATCTGC  720

Qy  727 CCCAAGGTGAACGCAGTGCCCCTGACCGGAGAAGATCAAGGCCCCATCTGC  786
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 CCCAAGGTGAACGCAGTGCCCCTGACCGGAGAAGATCAAGGCCCCATCTGC  780

Qy  787 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCGAGATCGGCGCGC  846
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781 GAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCGAGATCGGCGCGC  840

Qy  847 CCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGC  906
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db 1549 GCCCAGCTGGCGAGCTGGCCGAGACCGGCGAGATCTCGCCGAGATCCGCCGAGCCGTGTAC  1608
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1627 TACGACCCCAGCCAAGGACCACTGGTGCCGCCGAGATCCAGAGCAGGCAGGCCCACGACCAGTGACC  1686
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1609 TACGACCCCAGCCAAGGACCAAGGACCACTGGTGCCGCCGAGATCCAGAGCAGGCAGGCCCACGACCAGTGACC  1668

Qy 1687 TACCAGATCTACCAGGACCGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATCGCGC  1746
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1669 TACCAGATCTACCAGGACCGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATCGCGC  1728

Qy 1747 ACCGCCCACCAACGACGACGTGAAGCACGTGACCGAGCCGTGACCGAGGAAGATCGCCATGCAGAG  1806
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1729 ACCGCCCACCAACGACGACGTGAAGCACGTGACCGAGCCGTGACCGAGGAAGATCGCCATGCAGAG  1788

Qy 1807 AGCATCGTGATCTGGGCAAGACACCCCAAGTTCCGCCTCCCATCCAGAAGGAGACCCTGG  1866
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1789 AGCATCGTGATCTGGGCAAGACACCCCAAGTTCCGCCTCCCATCCAGAAGGAGACCCTGG  1848

Qy 1867 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC  1926
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1849 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC  1908

Qy 1927 ACCCCCCCTGTGAAGCTGGTACCAGCTGGAGAAGAGGAGCCCATCATCCGCGGCCGAG  1986
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1909 ACCCCCCCTGTGAAGCTGGTACCAGCTGGAGAAGAGGAGCCCATCATCCGCGGCCGAG  1968

Qy 1987 ACCTTCTACGTGGACGGCGCCGCCAACCGCGCCAAGCCGAGAGGCCGAGGCCCGCTACCTG  2046
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1969 ACCTTCTACGTGGACGGCGCCGCCAACCGCGCCAAGCCGAGAGGCCGAGGCCCGCTACCTG  2028

Qy 2047 ACCGACCGGGGCCCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG  2106
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2029 ACCGACCGGGGCCCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAG  2088

Qy 2107 CTGCAGGCCATCCAGCTGGGGCATCATCCAGGCCCTGACGGAGGCGCAGCGGAGCTGGTG  2166
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2089 CTGCAGGCCATCCAGCTGGGGCATCATCCAGGCCCTGACGGAGGCGCAGCGGAGCTGGTG  2148

Qy 2167 AGCCAGTAGCCCTGGGCATCATCCAGGCCCGACAAGAGCGAGAGCGAGAGCGAGCTGGTG  2226
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2149 AGCCAGTAGCCCTGGGCATCATCCAGGCCCGACAAGAGCGAGAGCGAGAGCGAGCTGGTG  2208

Qy 2227 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCCCC  2286
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2209 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCCGC  2268

Qy 2287 CACAAGGGCATCGGCGGCAACGAGCAGATCGAGCAAGCTGGTGAGCAAGGGCATCCGCAAG  2346
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2269 CACAAGGGCATCGGCGGCAACGAGCAGATCGAGCAAGCTGGTGAGCAAGGGCATCCGCAAG  2328

Qy 2347 GTGCTGTTCCTGGACGGCATCGACAAGGCGCAGGACGAGACCAGTACACATGGACGACCTG  2406
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2329 GTGCTGTTCCTGGACGGCATCGACAAGGCGCAGGACGAGACCAGTACACATGGACGACCTG  2388

Qy 2407 TACGTGGGCAGCGGCAGCGGCGGCGGGCCCTAGGATGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2463
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2389 TACGTGGGCAGCGGCAGCGGCGGCGGGCCCTAGGATGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2445
```

RESULT 6
US-10-190-305A-37
; Sequence 37, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05

```
Db   841 CCCGTGTTCGCCATCAAGAAGAAGGACGACCACCAAGTGGGCGAAGCTGGTGGACTTCCGC 900
Qy   907 GAGCTGAACAAGCGCACCCCAGGACTTCTCTGGGAGGTGCAGCTGCAGGGCATCCCCACCCCGCC 966
Db   901 GAGCTGAACAAGCGCACCCCAGGACTTCTGGGAGGTGCAGCTGCAGGGCATCCCCACCCCGCC 960

Qy   967 GGCCTGAAGAAGAAGAGAGCGTGACCGTGCTGCGACGTGGGCGACGGCACCGCCTACTTCAGCGTG 1026
Db   961 GGCCTGAAGAAGAAGAGAGCGTGACCGTGCTGCGACGTGGGCGACGGCACCGCCTACTTCAGCGTG 1020

Qy  1027 CCCCTGACGACGGACTTCCGCAAGTACACCCGCCTTCACCATCCCCAGCATCAACAACGAG 1086
Db  1021 CCCCTGGACGACGACTTCCGCAAGTACACCCGCCTTCACCATCCCCAGCATCAACAACGAG 1080

Qy  1087 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1146
Db  1081 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1140

Qy  1147 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1206
Db  1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATC 1200

Qy  1207 GTGATCTACCAGTACAATGGACGACCTGTACGTGGGCAGCGACCTGGGCCAGCAC 1266
Db  1201 GTGATCTACCA------GGCCCCCGTGTACGTGGGCAGCGACCTGGGCCAGCAC 1254

Qy  1267 CGCGCCCAAGATCGAGGAGCTGCGCCAAGCACCTGTCGCTGCGCTGGGGCTTCACCACCCCCGAC 1326
Db  1255 CCGCGCCCAAGATCGAGGAGCTGCGCCAAGCACCTGTCGCTGCGCTGGGGCTTCACCACCCCCGAC 1314

Qy  1327 AAGAAGACCACCAGAAGAGGAGCCCCCTTCCGTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1386
Db  1315 AAGAAGCACCAGAAGAGGAGCCCCCCTTCCT------CGAGCTGCACCCCGACAAG 1368

Qy  1387 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGACCGTGAACGACATCCAG 1446
Db  1369 TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGACCGTGAACGACATCCAG 1428

Qy  1447 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCAGCTCACCCCGGCATCAAGTGCGCCCAG 1506
Db  1429 AAGCTGGTGGGCAAGCTGAACTGGGCCAGCAGCTCACCCCGGCATCAAGTGCGCCCAG 1488

Qy  1507 CTGTGCAAGCTGCTGCGCGGCCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1566
Db  1489 CTGTGCAAGCTGCTGCGCGGCCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1548

Qy  1567 GCCGAGCTGGAGCTGGCCGAGAACCGGGCCGAGATCCTGCGCGGAGCCCGTGCACGCCGGTGTAC 1626
Db  1549 GCCGAGCTGGAGCTGGCCGAGAACCGGGCCGAGATCCTGCGCGGAGCCCGTGCACGCCGGTGTAC 1608

Qy  1627 TACGACCCCAGCAAGGACCTGGTTGCCGCAGATCCCGATCTCGTGCCCAGGGCCACGACCAGTGGACC 1686
Db  1609 TACGACCCCAGCAAGGACCTGGTTGCCGCAGATCCCGATCTCGTGCCCAGGGCCACGACCAGTGGACC 1668

Qy  1687 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1746
Db  1669 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1728

Qy  1747 ACCGCCCCACCACCACGACGTGAAGCAGCTGACGGCCGTGCTGCAGAGATCGGCCATGGAG 1806
Db  1729 ACCGCCCCACCACCACGACGTGAAGCAGCTGACGGCCGTGCTGCAGAGATCGGCCATGGAG 1788

Qy  1807 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCGGCATCCGAAGGAAGGACCTGG 1866
Db  1789 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCGTCCGCCTCCAGAAGGAAGGACCTGG 1848

Qy  1867 GAGAGCCTGGTGCACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1926
Db  1849 GAGAGCCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1908

Qy  1927 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCCGCGCCCAG 1986
Db  1909 ACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCCGCGCCGCGCAG 1968
```

```
Qy  1987 ACCTTCTACGTGACGGCCGGCGGCCGCCAACCGCGAGACCAAGATCGGCGAGACCGGCCTACGTG 2046
Db  1969 ACCTTCTACGTGACGGCCGGCGGCCGCCAACCGCGAGACCAAGATCGGCGAGACCGGCCTACGTG 2028

Qy  2047 ACGACCCGGGCCGGCCGGCCAGAGATCGTGAGCCTGACGCCTGACCGGAGACCAACAGAAGACCGAG 2106
Db  2029 ACGACCCGGGCCGGCCGGCCAGAGATCGTGAGCCTGACCGGAGACCAACAGAAGACCGAG 2088

Qy  2107 CTGCAGGCCATCAGCTGGCCTGGAGGACAGCGGCAGCGAGTGAACATCGTGACCGAC 2166
Db  2089 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGTGAACATCGTGACCGAC 2148

Qy  2167 AGCCAGTACGGCCCTGGGCATCATCCAGGCCCAGGCCCGCACAAGAGCGAGAGCGAGCTGGTG 2226
Db  2149 AGCCAGTACGGCCCTGGGCATCATCCAGGCCCAGGCCCGCACAAGAGCGAGAGCGAGCTGGTG 2208

Qy  2227 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2286
Db  2209 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCC 2268

Qy  2287 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2346
Db  2269 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2328

Qy  2347 GTGCTGTTCCTGGACGGCATCGATGGCCATCGATCCTGATCTCGAGTGGATCGATTAAAAGCTTCCCGGGCTAGCACCCGT 2463
Db  2329 GTGCTGTTCCTGGACGGCATCGATGGCCATCGATCCTGATCTCGAGTGGATCGATTAAAAGCTTCCCGGGCTAGCACCCGGT 2445

Qy  2407 TACGTGGGCAGCGACGGCGGCGGGCCCTAGGGATCGATTAAAAGCTTCCCGGGCTAGCACCCGT 2463
Db  2389 TACGTGGGCAGCGACGGCGGCGGGCCCTAGGGATCGATTAAAAGCTTCCCGGGCTAGCACCCGGT 2445
```

RESULT 7
US-10-190-435-9
; Sequence 9, Application US/10190435
; Publication No. US2003014324 8A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP1813.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmut_C
US-10-190-435-9

Query Match   97.0%;   Score 2394.8;   DB 14;   Length 3930;
Best Local Similarity 99.2%;   Pred. No. 0;
Matches 2431;   Conservative 0;   Mismatches 7;   Indels 12;   Gaps 2;

```
Qy    14 TCGGCCGGACCCATGAGCCATGAGCCAGGCCAGCCACCAACATCCTGATGCGCGCAGCAACTTCA 73
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1487 TCGCCGAGGCCATGAGCCATGAGCCAGGCCAGCCACCAACATCCTGATGCGCGCAGCAACTTCA 1546

Qy    74 AGGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATCGCGCCCGCA 133
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1547 AGGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCGCCCGCA 1606

Qy   134 ACTGCGCGCCCCCCGCCCCCGCAAGAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGGCACCAGATGA 193
         |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1607  ACTGCCCGCGCCCCCGCAAGAGGGCGTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA  1666
Qy   194  AGGACTGCACCGACGGCGCCAGGCGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA   253
Db  1667  AGGACTGCACCGACGGCGCCAGGCCAACTTCTTCCGCGAGGACCTTGGCCTTCCCCAGGGCA  1726
Qy   254  AGCCCGCGGAGTTCCCCAGCGACGAGACCGGCCAACAGCGCCAACAGCGCCAGCGCCGAGCTGC   313
Db  1727  AGCCCGCGGAGTTCCCCAGCGAGCGAGACCGGCCAACAGCGCCAACAGCGCCAGCGCCGAGCTGC  1786
Qy   314  AGGTGCGGCGGACAACCCCGCGAGCGCCGGCCGGCCGAGCGCCAGGGCCACCCTGAACT   373
Db  1787  AGGTGCGGCGGGACAACCCCGCGAGCGCCGGCCGGCCGAGCGCCAGGGCCACCCTGAACT  1846
Qy   374  TCCCCCAGATCACCCTGTGCGACGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA   433
Db  1847  TCCCCCAGATCACCCTGTGCGACGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA  1906
Qy   434  AGGAGGCCCTGCTGTGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG   493
Db  1907  AGGAGGCCCTGCTGTGACACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG  1966
Qy   494  GCAAGTGGAAGCCCAAGATGATGGGCGGCATCGGCGCGCTTCATCAAGGTGCGCCAGTACG   553
Db  1967  GCAAGTGGAAGCCCAAGATGATGGGCGGCATCGGCGCGCTTCATCAAGGTGCGCCAGTACG  2026
Qy   554  ACCAGATCCTGATCGAGATCTCGGCGCAAGAAGCCATCGGCACCCGTGCTGCTGATCGGCCCA   613
Db  2027  ACCAGATCCTGATCGAGATCTCGGCGCAAGAAGCCATCGGCACCGTGCTGCTGATCGGCCCCA  2086
Qy   614  CCCCGTGAACATCATCGGCCGACAACATGCTGACCCAGTCGGGTGCACCCTGAACTTCC   673
Db  2087  CCCCGTGAACATCATCGGCCGACAACATGCTGACCCAGTCGGGTGCACCCTGAACTTCC  2146
Qy   674  CCATCAGCGCCCATCGACGACCGTGAAGCTGAAGCCCGGCATGGACGGCCCCCAAGG   733
Db  2147  CCATCAGCGCCCATCGACGACCGTGAAGCTGAAGCCCGGCATGGACGGCCCCCAAGG  2206
Qy   734  TGAAGCAGTGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA   793
Db  2207  TGAAGCAGTGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA  2266
Qy   794  TGGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGT   853
Db  2267  TGGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGT  2326
Qy   854  TCGGCCATCAAGAAGGAGGACAGCAGGTGTGGTGGGACTTCCGCGACGAGCTGA   913
Db  2327  TCGGCCATCAAGAAGGAGGACAGCAGGTGTGGTGGGACTTCCGCGACGAGCTGA  2386
Qy   914  ACAAGCGCACCCAGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGGCCTGA   973
Db  2387  ACAAGCGCACCCAGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGGCCTGA  2446
Qy   974  AGAGAGAAGAGAGCGTGACCGTGCTGCTGGGACGTGGGCGACGGCACCCCGAGCGTGA  1033
Db  2447  AGAGAGAAGAGAGCGTGACCGTGCTGCTGGGACGTGGGCGACGGCACCCCGAGCGTGA  2506
Qy  1034  ACGAGGACTTCCGCAAGTACGCCGGCTTCACCCTTCCACCGCAGCGAGACCCCCG  1093
Db  2507  ACGAGGACTTCCGCAAGTACGCCGGCTTCACCCTTCCACCGCAGCGAGACCCCCG  2566
Qy  1094  GCATCCGGCTACCAACCGTCGTCTGCCCGCCGTCTGGGCTGAAGGCTGGAAGGGCGCCCCGAGCATCTTCC  1153
Db  2567  GCATCCGGCTACCAACCGTCGTCTGCCCGCCGTCTGGGCTGAAGGCTGGAAGGGCAGGGCCCCGAGCATCTTCC  2626
Qy  1154  AGAGCAGCATGACCAAGATCCTTGGAGCCTTCCGCGCCCGCAACCCCGAGATCGTGATCT  1213
Db  2627  AGAGCAGCATGACCAAGATCCTTGGAGCCTTCCGCGCCCGCAACCCCGAGATCGTGATCT  2686
Qy  1214  ACCAGTACATGACGACCTGTACGTGGGCAGCGACCTGAGCATCGGCCAGCACCGCGCCA  1273
Db  2687  ACCA------GGCCCCCGTACGTGGGCAGCGACCTGAGCATCGGCCAGCACCGCGCCA  2740

Qy  1274  AGATCGAGGAGCACTGCGCGGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAGAGC  1333
Db  2741  AGATCGAGGAGCTGCGGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAACAAGAGAGC  2800
Qy  1334  ACCAGAAGGAGAGCCCCCTTCCTGTGATGGCTACGAGGTGCACCCGACAAGTGGACCG  1393
Db  2801  ACCAGAAGGAGAGCCCCCCTTCTGTGATGGC-----CGAGCTGCACCCCGACAAGTGGACCG  2854
Qy  1394  TGCAGCCCATCGAGCTGCCCGAGAGAGAGCTGAACCGTGAACGACATCCAGAAGCTGG  1453
Db  2855  TGCAGCCCATCGAGCTGCCCGAGAGGAGAGCTGACCGTGAAACACATCCAGAAGCTGG  2914
Qy  1454  TGGGCAAGTGAACTGGGCCAGCAGATCTACCCCGGCATCAAGGTGCCCAGCTGTGCA  1513
Db  2915  TGGGCAAGTGAACTGGGCCAGCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA  2974
Qy  1514  AGCTGCTGCGCGGCGCTGAAGGCCGTGACCCTGACCGACCATCGTCGCCCTGACCGAGGAGGCCGAGC  1573
Db  2975  AGCTGCTGCGCGGCGCCCCAAGGCCGTGACCCTGACCGACCATCGTCGCCCTGACCGAGGAGGCCGAGC  3034
Qy  1574  TGGAGCTGGCCGAGAACCGCGAGATCTGCGCGAGAACCTGCGCCGGCCGTGCACCGCCGCGACC  1633
Db  3035  TGGAGCTGGCCGAGAACCGCGAGATCTGCGCGAGAACCTGCGCCGGCCGTGTACTACGACC  3094
Qy  1634  CCAGCAAGGACCTGGTTGCGGCGAGATCCAGAAGCACAGGGCCACGACCAGTGGACCTACCAGA  1693
Db  3095  CCAGCAAGGACCTGGTTGCGGCGAGATCCAGAAGCACAGGGCCACGACCAGTGGACCTACCAGA  3154
Qy  1694  TCTACCAGGAGGCCCTTCAAGAACCTGAAGACCTGCAGGCCAAGATGGCGCACCGCCC  1753
Db  3155  TCTACCAGGAGGCCCTTCAAGAACCTGAAGACCTGCAGGCCAAGATGGCGCACCGCCC  3214
Qy  1754  ACACCAACGAGCTGAAGCAGCTGACCGAGACGGCCGTGCCGAGAGATCGGCATGGAGAGGAGCATCG  1813
Db  3215  ACACCAACGAGCTGAAGCAGCTGACCGAGACGGCCGTGCCGAGAGATCGGCATGGAGAGGAGCATCG  3274
Qy  1814  TGATCTGGGGCAAGACCCCCAAGTTCCGGCCTGGGATCCCGAGTGGGAGTTCGTGGAACACCCCCCC  1873
Db  3275  TGATCTGGGGCAAGACCCCCAAGTTCCGGCCTGGGATCCCGAGTGGGAGTTCGTGGAACACCCCCCC  3334
Qy  1874  GGTGGACCGACTACTGGCAGGCCACCTGAGGAGCCCATCATCGGCGCGCGGCCTGACCTGTGACCGACC  1933
Db  3335  GGTGGACCGACTACTGGCAGGCCACCTGAGGAGCCCATCATCGGCGCGCGGCCTGACCTGTGACCGACC  3394
Qy  1934  CCCTGGTGAAGCTGTGTACCAGCTGAGAGAGGAGCCCAACCGGCCAAGATCGGCAAGGCCGGCTACGTGACCGACC  1993
Db  3395  CCCTGGTGAAGCTGTGTACCAGCTGAGAGAGGAGCCCAACCGGCCAAGATCGGCAAGGCCGGCTACGTGACCGACC  3454
Qy  1994  ACGTGGACGGCGCCGCGCAAGAGATCGTGACCGGCCTGAGCCTGACCAGCCAACCAGGAGCTGCCAGG  2053
Db  3455  ACGTGGACGGCGCCGCGCAAGAGATCGTGACCGGCCTGAGCCTGACCAGCCAACCAGGAGCTGCCAGG  3514
Qy  2054  GGGGCCCGGCGAGAAGATCGTGAGGCCTGAGCCCGGACCGAGCTGTGCAGGCCACCAGAGACCGAGCTGCAGG  2113
Db  3515  GGGGCCCGGCGAGAAGATCGTGAGGCCTGAGCCCGGACCGAGCTGTGCAGGCCACCAGAGACCGAGCTGCAGG  3574
Qy  2114  CCATCCAGCTGGCCCTGCCAGACAGCGGCAGCGGCACGGCAGCAGCCCGACAGCAGCCAGT  2173
Db  3575  CCATCCAGCTGGCCCTGCCAGACAGCGGCAGCGGCACGGCAGCAGCCCGACAGCAGCCAGT  3634
Qy  2174  ACGCCCTGGCATCCAGGCCAGACAAGAGCGAGCTGGTGAACCAGA  2233
Db  3635  ACGCCCTGGCATCCAGGCCAGACAAGAGCGAGCTGGTGAACCAGA  3694
Qy  2234  TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTTGGATCGT  2293
Db  3695  TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTTGGATCGT  3754
Qy  2294  GCATCGGCGGCGCAACGAGCAGATCGACAAGTGCTGTGAGGCAGGGCATCCGCCAAGGTGCGT  2353
Db  3755  GCATCGGCGGCGCAACGAGCAGATCGACAAGTGCTGTGAGGCAGGGCATCCGCCAAGGTGCGT  3814
```

```
Qy   2354  TCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG  2413
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3815  TCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG  3874

Qy   2414  GCAGCGGCGGCGGCCGCCTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT  2463
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3875  GCAGCGGCGGCGGCCGCCTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT  3924


RESULT 8
US-10-190-435-10
; Sequence 10, Application US/10190435
; Publication No. US2003014324BA1
GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutArt_C
US-10-190-435-10

  Query Match      96.9%;  Score 2393.2;  DB 14;  Length 3930;
  Best Local Similarity   99.2%;  Pred. No. 0;
  Matches 2430;  Conservative  0;  Mismatches  0;  Indels  12;  Gaps  2;

Qy     14  TGGCCGAGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA    73
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487  TCGCCGAGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA  1546

Qy     74  AGGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGGA   133
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547  AGGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCGCA  1606

Qy    134  ACTGCCGCGGCCCCCCGCGGAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA   193
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607  ACTGCCGCGGCCCCCCGCGGAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA  1666

Qy    194  AGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA   253
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667  AGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA  1726

Qy    254  AGGCCCGCGAGTTCCCAGGCAGAGCAACCGCGCAACAGCCCCACCAGCCCGCGGAGCTGC   313
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727  AGGCCCGCGAGTTCCCAGGCAGAGCAACCGCGCAACAGCCCCACCAGCCCGCGGAGCTGC  1786

Qy    314  AGGTGCGCGGCGACAACCCCGCAGCGGAGGCCGGCGCGGAGGCCGGCGCACCCTGAACT   373
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787  AGGTGCGCGGCGACAACCCCGCAGCGGAGGCCGGCGCGGAGGCCGGCGCACCCTGAACT  1846

Qy    374  TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGTGAGCATCAAGGTGGCGGCCAGATCA   433
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847  TCCCCCAGATCACCCTGTGGCAGCGGCCCCCTGTGAGCATCAAGGTGGCGGCCAGATCA  1906

Qy    434  AGGAGGGCCCTGCTGGACACCCGGCGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG   493
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907  AGGAGGGCCCTGCTGGACACCCGGCGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCG  1966

Qy    494  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCCTTCATCAAGGTGCGCCAGTACG   553
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCCTTCATCAAGGTGCGCCAGTACG  2026
```

```
Qy    554  ACCAGATCTGATCGAGATCTGCGCAAGAAGGCCATCGGCAAGGTGCACCGTGCTGATCGGCAGG   613
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027  ACCAGATCTGATCGAGATCTGCGCAAGAAGGCCATCGGCACCGTGCTGATCGGCAGG  2086

Qy    614  CCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGCTGCACCCTGAACTTCC   673
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2087  CCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGCTGCACCCTGAACTTCC  2146

Qy    674  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGGCATGGACGATGGACGGCCCAAGG   733
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2147  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGATGGACGGCCCAAGG  2206

Qy    734  TGAAGCAGTGCCCCTGACCGAGAGAAGATCAAGGCCCTGACGGCCATCTGCGAGGAGA   793
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2207  TGAAGCAGTGCCCCTGACCGAGAGAAGATCAAGGCCCTGACGGCCATCTGCGAGGAGA  2266

Qy    794  TGGGAAGGAGGGCAAGATCACCAAGATGGCCCCGAGAACCCCTACAACACCCCCGTGT   853
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2267  TGGGAAGGAGGGCAAGATCACCAAGATGGCCCCGAGAACCCCTACAACACCCCCGTGT  2326

Qy    854  TCGCCATCAAGAAGGACAGCAGCCACCAAGTGGCCGAAGCTTCCGGACTTCCGGAGCTGA   913
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327  TCGCCATCAAGAAGGACAGCAGCCACCAAGTGGCCGAAGCTTCCGGACTTCCGGAGCTGA  2386

Qy    914  ACAAGGCCACCCAGGACTTCTTGGAGGTGCAGCTGGGCATCCCCACCCCGGCCCTGA   973
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387  ACAAGGCCACCCAGGACTTCTTGGAGGTGCAGCTGGGCATCCCCACCCCGGCCCTGA  2446

Qy    974  AGAAGAAGAGCGTGACCGTGCTGGACGTGGGGCACGGCCTACTTCAGCGTGCCCCTGG  1033
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2447  AGAAGAAGAGCGTGACCGTGCTGGACGTGGGGCACGGCCTACTTCAGCGTGCCCCTGG  2506

Qy   1034  ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCGCATCAACAACAGAGACCCCG  1093
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2507  ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCGCATCAACAACAGAGACCCCG  2566

Qy   1094  GCATCCGCTACCAGTACAAGTGCTGCCCCAGGGTGGAAGGGCAGCCCCAGCATCTTCC  1153
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2567  GCATCCGCTACCAGTACAAGTGCTGCCCCAGGGTGGAAGGGCAGCCCCAGCATCTTCC  2626

Qy   1154  AGAGCAGCATGACCAAGATCCTGAGCGCCCGCCCGAACCCCGAGATCGTGATCT  1213
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627  AGAGCAGCATGACCAAGATCCTGAGCGCCCGCCCGAACCCCGAGATCGTGATCT  2686

Qy   1214  ACCAGTACATGACGACGACCTGTACGTGGCGCAGCGGCCAGCACCGCGCCA  1273
                 ||||||||||||||||||   ||||||||||||||||||||||||||
Db   2687  ACCA------GGCCCCCCGTACGTGGCGCAGCGGCCAGCACCGCGCCA  2740

Qy   1274  AGATCGAGGAGCTGCGCAAGCACCTGCTGCGCGCTGGGCTTCACCACCCCGACAAGAAGC  1333
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741  AGATCGAGGAGCTGCGCAAGCACCTGCTGCGCGCTGGGCTTCACCACCCCGACAAGAAGC  2800

Qy   1334  ACCAGAGGAGCCCCTTCTGTGTGGATGGGCTACGAGCTGCACGCAGAGTGACCG  1393
                 ||||||||||||||||||||||||              |||||||||||||||
Db   2801  ACCAGAGGAGCCCCTTCTGTGTGGAT------CGAGCTGCACCCGACAAGAAGTGGACCG  2854

Qy   1394  TGGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAACTGG  1453
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2855  TGGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAACTGG  2914

Qy   1454  TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCCATCAAGGTGTGGCGCCAGATCGCA  1513
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2915  TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCCATCAAGGTGTGGCGCCAGATCGCA  2974

Qy   1514  AGTGCTGCGCGGCCCAAGGCCCTGACCGTGCCCCTGAGCGAGGAGGCCGAGC  1573
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2975  AGTGCTGCGCGGCCCAAGGCCCTGACCGTGCCCCTGAGCGAGGAGGCCGAGC  3034

Qy   1574  TGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGGAACTCATCAAGGTGCGCCAGTACACCGACC  1633
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3035  TGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGGAACTCATCAAGGTGCGCCAGTACACCGACC  3094
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3930
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagComplPolmutina_C
US-10-190-435-11

Query Match          96.9%;     Score 2393.2;    DB 14;    Length 3930;
Best Local Similarity 99.2%;     Pred. No. 0;
Matches 2430;      Conservative   0;  Mismatches   8;   Indels  12;   Gaps   2;
```

```
Qy      14 TGGCCGAGGCCATGAGCCAGCCAGCCGGCCACAATCCTGATGCAGCAGCAACTTCA    73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1487 TCGCCGAGGCCATGAGCCAGCCAGCCGGCCACAATCCTGATGCAGCAGCAACTTCA  1546

Qy      74 AGGGCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA   133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1547 AGGGCCCAAGGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA  1606

Qy     134 ACTGCCGCGCCCCCCGCCGAAGAAGAAGGGCTGCGGAAGTGCGGCAAGGAGGGCCACCAGATGA   193
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1607 ACTGCCGCGCCCCCCGCCGAAGAAGAAGGGCTGCGGAAGTGCGGCAAGGAGGGCCACCAGATGA  1666

Qy     194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGAACCGGCCTTCCCCCAGGGCA   253
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1667 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGAACCGGCCTTCCCCCAGGGCA  1726

Qy     254 AGGCCCGCGAGTTCCCCAGCCAGCAGAACCGGCGCCAACAGCCCCGCGAGCGGCCACCCTGAACTGC   313
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1727 AGGCCCGCGAGTTCCCCAGCCAGCAGAACCGGCGCCAACAGCCCCGCGAGCGGCCACCCTGAACTGC  1786

Qy     314 AGGTGCGGCGGCAACCCCGACAACCACCACCCCCGCCAGCGGCCACCCTGAACT   373
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1787 AGGTGCGGCGGCAACCCCGACAACCACCACCCCCGCCAGCGGCCACCCTGAACT  1846

Qy     374 TCCCCCAGATCACCCTGGCAGCGCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA   433
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1847 TCCCCCAGATCACCCTGGCAGCGCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA  1906

Qy     434 AGGAGGCCCTGCTGGACACGAACACCGTGCGGCATCGGCGGCATCGGCGTGCTGCGCCG   493
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1907 AGGAGGCCCTGCTGGACACGAACACCGTGCGGCATCGGCGGCATCGGCGTGCTGCGCCG  1966

Qy     494 GCAAGTGGAAGCCAAGATGATCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG   553
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1967 GCAAGTGGAAGCCAAGATGATCGGCATCGGCGGCTTCATCAAGGTGCGGCCAGTACG  2026

Qy     554 ACCAGATCCTGATCGAGATCTGCGGCAAGCATCGGCATCCGCACCGGCTGGGCCCCA   613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2027 ACCAGATCCTGATCGAGATCTGCGGCAAGCATCGGCATCCGCACCGGCTGGGCCCCA  2086

Qy     614 CCCCCGTGAACATCATCGAGAACCGTGCCCGTGAAGCTGAAGCTGAAGCAGCGGCCCAAGG   673
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2087 CCCCCGTGAACATCATCGAGAACCGTGCCCGTGAAGCTGAAGCTGAAGCAGCGGCCCAAGG  2146

Qy     674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCTGGACGGACGGCCCAAGG   733
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2147 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCTGGACGGACGGCCCAAGG  2206

Qy     734 TGAAGCAGTGCCCCGACCGAGAGAAGATCAAGGCCCTGACCGCCATCTGCGGAGAGA   793
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2207 TGAAGCAGTGCCCCGACCGAGAGAAGATCAAGGCCCTGACCGCCATCTGCGGAGAGA  2266

Qy     794 TGGGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGT   853
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2267 TGGGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCGTGT  2326

Qy     854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGGCCAAGCTGGTGGACTTCCGGCAGCTGA   913
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2327 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGGCCAAGCTGGTGGACTTCCGGCAGCTGA  2386
```

```
Qy    1634 CCAGCAAGGACCTGTGTGGCCGAGATCCAGGAGCAGGGCCACGACCAGTGACCTACCAGA  1693
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3095 CCAGCAAGGACCTGTGTGGCCGAGATCCAGGAGCAGGGCCACGACCAGTGACCTACCAGA  3154

Qy    1694 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCGGCCAGTACGGCCAAGATGCCGACCGCCCC  1753
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3155 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCGGCCAGTACGGCCAAGATGCCGACCGCCCC  3214

Qy    1754 ACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCG  1813
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3215 ACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCG  3274

Qy    1814 TGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCT  1873
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3275 TGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCT  3334

Qy    1874 GGTGGACCGACTACTGCAGGCCACCTGGACTGGGAGTTCGTGAACACCCCCCA  1933
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3335 GGTGGACCGACTACTGCAGGCCACCTGGACTGGGAGTTCGTGAACACCCCCCA  3394

Qy    1934 CCCTGTGAAGCTGTGTGTACCAGCTGGAGAGGAGAGCCCATCATCGCGCCGAGACCTTCT  1993
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3395 CCCTGTGAAGCTGTGTGTACCAGCTGGAGAGGAGAGCCCATCATCGCGCCGAGACCTTCT  3454

Qy    1994 ACGTGACGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCGGCTACGTGACCGACC  2053
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3455 ACGTGACGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCGGCTACGTGACCGACC  3514

Qy    2054 GGGGCCGCGGCAGAAGATCGTGAGCCCTGCAGGACAGCGGCGCAGCGAGTGAACATCGTGACCGAGCAGT  2113
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3515 GGGGCCGCGGCAGAAGATCGTGAGCCCTGCAGGACAGCGGCGCAGCGAGTGAACATCGTGACCGAGCAGT  3574

Qy    2114 CCATCCAGCTGGCCCTGCCGGAGATCATCCAGGCCGAGGCGCAGCCCCAGCAGAGCTGTGACCAGA  2173
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3575 CCATCCAGCTGGCCCTGCCGGAGATCATCCAGGCCGAGGCGCAGCCCCAGCAGAGCTGTGACCAGT  3634

Qy    2174 ACGCCCTGGCATCATCCAGCCGAGCGCCCAGCCCCAGACAAGAGGAGCTCGGTGAACCAGA  2233
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3635 ACGCCCTGGCATCATCCAGCCGAGCGCCCAGCCCCAGACAAGAGCAGCTCGGTGAACCAGA  3694

Qy    2234 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGAGCTGGTGCCCCGCCCACAAGG  2293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3695 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGGAGCTGGTGCCCCGCCCACAAGG  3754

Qy    2294 GCATCCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCGAAGGTGCTGT  2353
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3755 GCATCCGGCGGCAACGACGACAGATCGACAAGCTGGTGAGCAAGGGCATCCGCGAAGGTGCTGT  3814

Qy    2354 TCCTGAACTACGACGCATCGATGATGCGGCCATCGTGATCTACCAGTACACTGAACGACACCTGTACGTGG  2413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3815 TCCTGAACTACGACGCATCGATGATGCGGCCATCGTGATCTACCAGTACACTGAACGACACCTGTACGTGG  3874

Qy    2414 GCAGCGGCGGCGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGGTAGCACCCGGT  2463
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3875 GCAGCGGCGGCGGCCCTAGGATCGATTAAAAAGCTTCCCGGGGGTAGCACCCGGT  3924
```

```
RESULT 10
US-10-190-435-58
; Sequence 58, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 58
;  LENGTH: 5184
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: TatRevNefgagCpolina C
US-10-190-435-58

Query Match      96.9%;  Score 2393.2;  DB 14;  Length 5184;
Best Local Similarity  99.2%;  Pred. No. 0;
Matches 2430;  Conservative  0;  Mismatches  0;  Indels  12;  Gaps  2;
```

```
Qy     14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACTTCA 73
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACTTCA 2800

Qy     74 AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 2860

Qy    134 ACTGCCGCGCGCCCCCGCAAGAGAGGGCTGCTGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 ACTGCCGCGCGCCCCCGCAAGAGAGGGCTGCTGAAGTGCGGCAAGGAGGGCCACCAGATGA 2920
```

```
Qy    194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCGGCGAGGACCTGGCCTTCCCCCAGGGCA 253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCGGCGAGGACCTGGCCTTCCCCCAGGGCA 2980

Qy    254 AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCGCCACCAGCGCGGAGCTGC 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCGCCACCAGCGCGGAGCTGC 3040

Qy    314 AGGTGCGCGGCGACAACCCCCGCAGCGAGCGCGGCCGGCGCCGAGCGCGACCCTGAACT 373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGCGGCGACAACCCCCGCAGCGAGCGCGGCCGGCGCCGAGCGCGACCCTGAACT 3100

Qy    374 TCCCCCAGATCACCCTGTGGCGAGCGCCCCTGGTGAGCAGCATCAAGGTGGGCGGCCAGATCA 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCCAGATCACCCTGTGGCGAGCGCCCCTGGTGAGCAGCATCAAGGTGGGCGGCCAGATCA 3160

Qy    434 AGGAGGCCCTGCTGACACCGGCGCCGACGACACCGTGCTCGGAGGAGATGAGCGTGCCCG 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGCCCTGCTGACACCGGCGCCGACGACACCGTGCTCGGAGGAGATGAGCGTGCCCG 3220

Qy    494 GCAGTGGAAGCCCAAGATGATCGGGGCATCGGCGCCTTCATCGGCGTGGCCGCCAGTACG 553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATCGGGGCATCGGCGCCTTCATCAAGGTGCCCAGTACG 3280

Qy    554 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGCCCCCA 613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCA 3340

Qy    614 CCCCCGTGAACATCATCGGCGCCAACATGCTGGCCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCCGTGAACATCATCGGCGCCAACATGCTGGCCTGACCCAGCTGGGCTGCACCCTGAACTTCC 3400

Qy    674 CCATCAGCCCCATCGAGACCCCGTGCCCGTGGAGCTGAAGCTGAAGCCCATGGACGGCCCCAAGG 733
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAGACCCCGTGCCCGTGGAGCTGAAGCTGAAGCCCATGGACGGCCCCAAGG 3460

Qy    734 TGAAGCAGTGGCCCCTGACCGGCGCTGGGAGGAGATCAAGGCCTTGACCGCCATCTGCAGGAGA 793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGGCCCCTGACCGGCGCTGGGAGGAGATCAAGGCCTTGACCGCCATCTGCAGGAGA 3520

Qy    794 TGGAGAAGGAGGGCCAAGATCACCAGCAGCACCAAGATCACCGGCGCCCCGAGAACCCCTACAACACCCCCGTGT 853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGAGAAGGAGGGCCAAGATCACCAGCAGCACCAAGATCACCGGCGCCCCGAGAACCCCTACAACACCCCCGTGT 3580

Qy    854 TCGCCATCAAGAGAAGGAGGGCTGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCAAGAGAAGGAGGGCTGCGCAAGCTGGTGGACTTCCGCGAGCTGA 3640

Qy    914 ACAAGGCGCACCCAGGACTTCTGGCAGTGCAGCTGGCATCCCCGCCGCCGCCTGA 973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGGCGCACCCAGGACTTCTGGCAGTGCAGCTGGCATCCCCGCCGCCCCGGCCTGA 3700

Qy    974 AGAAGAAGAGAGCGGTGACGCTGCTGACGTGCGTCTGGACGTGGGCGACGCCTACTTCAGCGGTGCCCCTGG 1033
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAAGAAGAGAGCGGTGACGCTGCTGACGTGCGTCTGGACGTGGGCGACGCCTACTTCAGCGGTGCCCCTGG 3760

Qy   1034 ACCAGGACATGACCAAGATCCTGGAGCGCCTTCACCATCCCAGCATCAACGAGCACCCCCCG 1093
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACCAGGACATGACCAAGATCCTGGAGCGCCTTCACCATCCCAGCATCAACGAGCACCCCCCG 3820

Qy   1094 GCATCCGCTACCAGTACAAACGTGCTGGCAGCGAGGGCAGCAGCCAGGGCATCTTCC 1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGCTACCAGTACAAACGTGCTGGCAGCGAGGGCAGCAGCCAGGGCATCTTCC 3880

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCGCCTTCGCGCCCTGCTGTGATCT 1213
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3881 AGAGCAGCATGACCAAGATCCTGGAGCGCCTTCGCGCCCTGCTGTGATCT 3940

Qy   1214 ACCAGTACATGACCAGTGTACGTGGCAGCGACCTGACGCGCACCGGCCA 1273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3941 ACCA------GGCCCCCGTGTACGTGGGCACGGCCGAGCACCTGGCCGCCCCA 3994

Qy   1274 AGATCGAGGAGGAGCTGCGCAAGCACCTGCTGCTCGGCTGCTGGTGGGCTTCACCACCCCGACAAGAAGC 1333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3995 AGATCGAGGAGGAGCTGCGCAAGCACCTGCTGCTCGGCTGCTGGTGGGCTTCACCACCCCGACAAGAAGC 4054

Qy   1334 ACCAGAAGGAGAGCGAGCCCCTTCTTGTGGATGGGCTTACGGCTGCTGGTGGCGCTTCACCACCCCGACAAGTGGACCG 1393
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4055 ACCAGAAGGAGAGCGAGCCCCTTCTTCCGCCCAT------CGAGCTGCACCCCGACAAGTGGACCG 4108

Qy   1394 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCCTGGACCGGTGAACGCATCGAGAACGACTTGCGGAACTGG 1453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4109 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCCTGGACCGTGAACGACATCAGGAACGTGG 4168

Qy   1454 TGGGCAAGCTGAACTGGGCCCAGCCGCGGCATCCCGGCATCGCCGCAGCTGTGTGCA 1513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4169 TGGGCAAGCTGAACTGGGCCCAGCCGCGGCATCCCGGCATCGGCCAGCTGTGTGCA 4228

Qy   1514 AGCTGCTGCGGCGCGGCCAAGGCCCTGACCGACATGTGCCCCCTGACCGGCGAGGAGGCCGAGC 1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4229 AGCTGCTGCGGCGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGGCGAGGAGGCCGAGC 4288

Qy   1574 TGGAGCTTGCCGAGAACCGCGAGATCCTGCGGCAGCCCGAGCCCGTGACGCGGCGTGTACTACGACC 1633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4289 TGGAGCTTGCCGAGAACCGCGAGATCCTGCGGCAGCCCGAGCCCGTGACGCGGCGTGTACTACGACC 4348

Qy   1634 CCAGCAAGAACCTGGTGCCGAGAACCAGCAGGGCCACAGCAGGCCACCAGTGGACCTACCAGA 1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4349 CCAGCAAGAACCTGGTGCCGAGAACCAGCAGGGCCACAGCAGGCCACCAGTGGACCTACCAGA 4408

Qy   1694 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCTACCGGCAAGTACCGCCTGCCCCAAGTTCGCCCATCGAGACCGGGACCT 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4409 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCTACCGGCAAGTACCGCCTGCCCCAAGTTCGCCCATCGAGACCGGGACCT 4468

Qy   1754 ACACCAACGACAGTGAAGCAGCTGACCGACCCGAGCCCGTGCAGAAGATCGCAGAGAGCATCG 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4469 ACACCAACGACAGTGAAGCAGCTGACCGACCCGAGCCCGTGCAGAAGATCGCAGAGAGCATCG 4528

Qy   1814 TGATCTGGGCAAGACCCCATCGGCCCTGACCGTCCCAAGTTCGCCCATCGAGAGGAGACCT 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4529 TGATCTGGGCAAGACCCCATCGGCCCTGACCGTCCCAAGTTCGCCCATCGAGAGGAGACCT 4588

Qy   1874 GGTGACCGACTACTGGCAGGCCCGGGTACCCGAGTGGAAGTTCGTGAACACCCCCC 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4589 GGTGACCGACTACTGGCAGGCCCACCTGGGGTGGAAGTTCGTGAACACCCCCC 4648

Qy   1934 CCCTGGTGAAGCTGTGGTTACCAAGTGGTTCCCGCGCCCATCATCGGCGCCGAGACCTTCT 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4649 CCCTGGTGAAGCTGTGGTTACCAGTGGTTCCCGCGCCCATCATCGGCGCCGAGACCTTCT 4708

Qy   1994 ACGTGGACGCGGCGCCCCAACCGCGCAGAAATCGTTGAGCGCCAACCAGAAGCTGACCGCAGG 2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4709 ACGTGGACGCGGCGCCCCAACCGCGCAGAAATCGTTGAGCGCCAACCAGAAGCTGACCGCAGG 4768

Qy   2054 GGGCGCCGGCAGAAGATCGTGAGCGCCCTGACGCGGCGACAGCGGCGAGAGGCTGACCCACC 2113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4769 GGGGCGCCGGCAGAAGATCGTGAGCGCCCTGACGCGGCGACAGCGGCGAGAGGCTGACCCACC 4828

Qy   2114 CCATCCGTGGCCCTGCGAGCGGCGGAGGGCATCCCAGCATCTGACAGCCAGT 2173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4829 CCATCCAGTGGCCCTGCGAGGCGGACAGCGGCGGAGTGAACATCGTGACCAGCCAGT 4888

Qy   2174 ACGGCCTGGGCATCAGTCCAGGGCCAGGGGCAGCCCAGCCGGTGGTGAACCAGA 2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4889 ACGGCCCTGGGCATCAGTCCAGGGGCAGCGGGGGCAGCGAGGAGAGCCAGACGA 4948

Qy   2234 TCATCGAGCAGCTGATCAAGAAGAAGGTGTTACCGAGCTGGTGCCCGCCACAAGG 2293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4949 TCATCGAGCAGCTGATCAAGAAGAAGGTGTTACCTGAGCTGGTGCCCGCCCACAAGG 5008

Qy   2294 GCATCGGCGGCAAGCGACCTGTACGGTGGGCAGGGCATCCGGCAAGGTGCTGT 2353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5009 GCATCGGCGGCAAGCGACCTGTACGGTGGGCAGGGCATCCGGCAAGGTGCTGT 5068

Qy   2354 TCCTGACGGCGATCGATGGCGCATCGTGATCTACACCAGTACCGGACGACCTGTACGTGG 2413
```

RESULT 11
US-10-190-305A-82
; Sequence 82, Application US/10190305A
; Publication No. US20030198621A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan
; APPLICANT: LIAN, Ying
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR
; TITLE OF INVENTION: TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 2302-18702 / 18702.002
; CURRENT APPLICATION NUMBER: US/10/190,305A
; CURRENT FILING DATE: 2002-07-05
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 82
; LENGTH: 5184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: TatRevNefgagCpolina C
US-10-190-305A-82

Query Match   96.9%;   Score 2393.2;   DB 14;   Length 5184;
Best Local Similarity 99.2%;   Pred. No. 0;
Matches 2430; Conservative 0; Mismatches 8; Indels 12; Gaps 2;

```
Qy     14  TCGCCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA   73
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741  TCGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA  2800

Qy     74  AGGGCCCCAAGGCCACCAAGGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCA   133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801  AGGGCCCCAAGGCCACCAAGGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCA  2860

Qy    134  ACTGCCGCGCCCCCGCGCAAGAGAGGGGTGCTGAAGTGCGGCAAGAGGGGCCACCAGATGA   193
           |||     ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861  ACTGCCGCGCCCCCGCGCAAGAGAGGGGTGCTGAAGTGCGGCAAGAGGGGCCACCAGATGA  2920

Qy    194  AGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA   253
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   2921  AGGACTGACCGAGCGGCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA  2980

Qy    254  AGGCCCCGCGAGTTCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGAGCTGC   313
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981  AGGCCCCGCGAGTTCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGCGAGCTGC  3040

Qy    314  AGGTGCGCGGACAACCCCCGCCAGCCGAGGCGCCAGGCGCCAGGGCGCCAGCACCCTGAACT   373
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041  AGGTGCGCGGACAACCCCCGCCAGCCGAGGCGCCAGGCGCCAGGGCGCCAGCACCCTGAACT  3100

Qy    374  TCCCCAGATCACCCTGTGGCAGCGCCCCCTGTGAGCATCAAGGTGGCGGCCAGATCA   433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101  TCCCCAGATCACCCTGTGGCAGCGCCCCCTGTGAGCATCAAGGTGGCGGCCAGATCA  3160

Qy    434  AGGAGGCCTGCTGGACACCGAGCCGGCCGGACGACACCGTGCTGGAGAGATAGAGCTGCGCCG   493
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161  AGGAGGCCTGCTGGACACCGAGCCGGCCGGACGACACCGTGCTGGAGAGATGAGCTGCGCCG  3220

Qy    494  GCAAGTGGAAGCCAAGATGATCGGCGGCATCGGCCGCTTCATCAAGGTGCGCCAGTACG   553
           |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   3221  GCAAGTGGAAGCCAAGATGATCGGCGGCATCGGCCGCTTCATCAAGGTGCGCCAGTACG  3280

Qy    554  ACCAGATCCTGATCGAGATCTGGCCGCACCGTGCTGCTGATCGGCCCCA   613
           |||||||||||||||||||||||||||||||||||||||||||||
Db   3281  ACCAGATCCTGATGCAGAGATCTGCGGCAAGAGAGGCCATCGGCAGCGTGATCGGCCCCA  3340

Qy    614  CCCCGTGAACATCATCGGCCGCAACATGCTGCACCGGCTGGGCTGCACCCTGAACTTCC   673
           ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   3341  CCCCGTGAACATCATCGGCCGCAACATCATCGGCCGCTGACCCAGCTGGGCTGCACCCTGAACTTCC  3400

Qy    674  CCATCAGCCCCATCGAGACCGTGCCGTGAAGTCGTGAAGCTGGGCATGAGCAGCGGCCCCAAGG   733
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401  CCATCAGCCCCATCGAGACCGTGCCGTGAAGTCGTGAAGCTGAAGCCCGCATGACGACGCCCCAAGG  3460

Qy    734  TGAAGCAGTGCCCCTGACCGAGGAAGATCAGGCCCTGACCGCCATCTGCGAGGAGA   793
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   3461  TGAAGCAGTGCCCCTGACCGAGGAAGATCAGGAGATCAAGGCCCCTGACCCCTGACCGCCATCTGCGAGGAGA  3520

Qy    794  TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT   853
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521  TGGGAGGAGGGGGCAAGATCACCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT  3580

Qy    854  TCGCCATCAAGAAGGACACCAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA   913
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   3581  TCGCCATCAAGAAGGACACCAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA  3640

Qy    914  ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA   973
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641  ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA  3700

Qy    974  AGAAGAAGAAGAGCCGTGACCGTGCTGGACGTGGACGTGGGCGACGCGCTACTTCAGCGTGCCCCTG  1033
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701  AGAAGAAGAAGAGCCGTGACCGTGCTGGACGTGGACGTGGGCGACGCGCTACTTCAGCGTGCCCCTGG  3760

Qy   1034  ACGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACAGACCCCCCG  1093
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761  ACGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACAGACCCCCCG  3820

Qy   1094  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC  1153
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821  GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC  3880

Qy   1154  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATCGTGATCT  1213
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3881  AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCCGAGATCGTGATCT  3940

Qy   1214  ACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCA  1273
           ||||                ||||||||||||||||||||||||||||||||||||||||
Db   3941  ACCA-----------GGCCCCCGTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCA  3994

Qy   1274  AGATCGAGAGCTGCGCAAGCACCTGCTGCTGTCTGGGCTTCACCACCCCCGACAAGAAGC  1333
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3995  AGATCGAGAGCTGCGCAAGCACCTGCTGCTGTCTGGGCTTCACCACCCCCGACAAGAAGC  4054

Qy   1334  ACCAGAGAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGACCAAGTGGACCG  1393
           ||||||||||||||||||||||||||||         ||||||||||||||||||||||
Db   4055  ACCAGAAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGACCAAGTGGACCG  4108

Qy   1394  TGCAGCCCATCGAGCTGCCCGAGAAGGAGGGCAAGGCGGCCGGCACCCTGAACT   1453
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4109  TGCAGCCCATCGAGCTGCCCGAGAAGGAGGGCAAGGCGGCCGGCACCCTGAACT  4168

Qy   1454  TGGGCAAGTGAACTGGGCGCCAGCATCACCGGGCGGCCAGATCA   1513
           ||||||||||||||||||||||||||||||||||||||||||
Db   4169  TGGGCAAGTGAACTGGGCGCCAGCATCACCGGGCGGCCAGATCA  4228

Qy   1514  AGTGTGCTGGTGGCACCGAGATCGTGCCCCTGACCGAGGAGGCCGAGC   1573
           ||||||||||||||||||||||||||||||||||||||||||||
Db   4229  AGTGTGCTGGTGGCACCGAGATCGTGCCCCTGACCGAGGAGGCCGAGC  4288

Qy   1574  TGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCGTGTACTACGACC  1633
           ||||||||||||||||||||||||||||||||||||||||||||
Db   4289  TGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGAGCCGTGTACTACGACC  4348

Qy   1634  CCAGCAAGGACCTGTGGCCGAGATCCAGAAGCAGGGCCAGAGCTGTGCA  1693
           |||||||||||| |||||||||||||||||||||||||||||||||
Db   4349  CCAGCAAGGACCTGTGGCCGAGATCCAGAAGCAGGGCCAGAGCTGTGCA  4408

Db   5069  TCCTGACGACGGCATCGATGCGGCGCATCGTGATCTACCAGTACAGCAATGAGACGACCTGTACGTGG  5128

Qy   2414  GCAGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGCTGCTAGCACCGGT  2463
           |||||||||||||||||||||||||||||||||||||||||
Db   5129  GCAGCGGCGGGGGCCCTAGGATCGATTAAAAGCTTCCCGGGCTGCTAGCACCGGT  5178
```

```
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence: GagPolmut_C
US-10-190-435-13

  Query Match    95.7%;  Score 2362.8;  DB 14;  Length 3531;
  Best Local Similarity  98.4%;  Pred. No. 0;
  Matches 2411;  Mismatches  0;  Conservative  27;  Indels  12;  Gaps  2;
```

```
Qy      14  TGGCCGAGGCCATGAGCCAGGCCACCGACCAACATCCTGATGCAGCGAGCAACTTCA     73
            ||||||||||  ||||||||||||||    |||    ||| ||||||||||||||||
Db    1088  TGGCCGAGGCGATGAGCCAGGCCAGCCAACCAGCGGTGATGATGCAGCGAGCAACTTTAAAA 1147

Qy      74  AGGGCCCCAAGGCCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCCACATCGCCCGCA    133
            ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    1148  AGGGCCCCAAGGCCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA   1207

Qy     134  ACTGCGCGGCCCCCCGAAGAGAGGGCTGCTGGAAGTGCGGCAAGGAGTGCGGCACCAGATGA 193
            |||||||  |||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    1208  ACTGCGCGGCCCCCGAAGAGAGGGCTGCTGGAAGTGCGGCAAGGAGTGCGGCACCAGATGA 1267

Qy     194  AGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCA 253
            ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    1268  AGGACTGCACCGAGCGGCGCCAACTTCTTCCGGGAGGACCTGGCCTTCCCCCAGGGCA   1327

Qy     254  AGGCCCGGGAGTTCCCAGCGGCCAACAGGACCGGCCAACAGCCCGCCCAGGGCCGAGCTGC 313
            ||||||||||  ||||||||  ||||||||||||||||||||||||||||||||||||||
Db    1328  AGGCCCGGGAGTTCCCAGCGGCGCCAACAGCCCGCCCAGGGCCGAGCTGC  1387

Qy     314  AGGTGCGGGCGACAACCCCGCAGGGAGGCCGGCCGCAGGGCCACCCTGAACT    373
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1388  AGGTGCGGGCGACAACCCCGCAGGGAGGCCGGCCGCAGGGCCACCCTGAACT    1447

Qy     374  TCCCCAGATCACCCGTGGCAGCGCCCCTGTGAGCATCAAGGTGCGGGCAGATCA  433
            ||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db    1448  TCCCCAGATCACCCGTGGCAGCGCCCCTGTGAGCATCAAGGTGGGCGGCCAGATCA 1507

Qy     434  AGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGATGAGCCTGCCCG  493
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1508  AGGAGGCCCTGCTGGACACCGGCGCCGACGACACCGTGCTGGAGATGAGCCTGCCCG 1567

Qy     494  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1568  GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 1627

Qy     554  ACCAGATCCTGATCGAGATCTGCGGCCAGAAGGCCATCGGCACCGTGCTGGTCGGCCCCA 613
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1628  ACCAGATCCTGATCGAGATCTGCGGCCAGAAGGCCATCGGCACCGTGCTGGTCGGCCCCA 1687

Qy     614  CCCCGTGACATCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 673
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1688  CCCCGTGACATCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 1747

Qy     674  CCATCAGCCCCATCGAGACGTGCCCGTGAAGTGAAGCCGGCATGACGGGCCCCAAGG      733
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1748  CCATCAGCCCCATCGAGACGTGCCCGTGAAGTGAAGCCGGCATGGACGGCCCCAAGG     1807

Qy     734  TGAAGCAGTGCCCCCTGACCGAGAAGATCAAGGCCCTGACCGCCATCTGCGGAGGAGA    793
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1808  TGAAGCAGTGCCCCCTGACCGAGAAGATCAAGGCCCTGACCGCCATCTGCGGAGGAGA   1867

Qy     794  TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 853
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1868  TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 1927

Qy     854  TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1928  TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 1987

Qy     914  ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 973
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1988  ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 2047
```

```
Qy    1694  TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACCGGCAAGATGCGCACCGCCC 1753
            |||||||||||||||||||||||||||||||||||||||  |||||||||||||||||||
Db    4409  TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGGCAAGATGCGCACCGCCC  4468

Qy    1754  ACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATGCCATGGAGAGCATCG   1813
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4469  ACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATGCCATGGAGAGCATCG   4528

Qy    1814  TGATCTGGGCCAAGACCCCCAAGTTCCGCCTGCCATCCAGAAGGAGACCTGGGAGACCT   1873
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4529  TGATCTGGGCCAAGACCCCCAAGTTCCGCCTGCCATCCAGAAGGAGACCTGGGAGACCT   4588

Qy    1874  GGTGCGACTACTGGCAGGCCACCTGGATCCCCGAGTTGGAGTTCGTGAACACCCCCC     1933
            |||| ||||||||||||||||||||||||||||||||||| |||||||||  |||||
Db    4589  GGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTTGGGAGTTCGTGAACAACCCCC  4648

Qy    1934  CCCTGGTGAAGCTGTGTGTACCAGTGGAGAAGAGCCCATCATCGGCGGCCGAGACCTTCT  1993
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    4649  CCCTGGTGAAGCTGTGTTACCAGTGGAGAAGAGCCCATCATCGGCGGCCGAGACCTTCT  4708

Qy    1994  ACGTGACGGCGCGCCAACCCGGAGACCAAGATCGGCAAGCCGGCCTACGTGACCGACC   2053
            ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    4709  ACGTGACGGCGCGCCAACCCGGAGACCAAGATCGGCAAGCCGGCTACGTGACCGACC   4768

Qy    2054  GGGGCCGGCAGAAGATCGTGAGCCTGAGCGCAGCGGGCAGCACCAGAGACCGAGCTGCAGG 2113
            ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    4769  GGGGCCGGCAGAAGATCGTGAGCCTGAGCGCACGCGAGCACCAGAGACCGAGCTGCAGG  4828

Qy    2114  CCATCCAGCTGCCCTGCAGACAGCGGCAGCGAGTGAACATCGTGACCGAGCAGCCAGT    2173
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4829  CCATCCAGCTGCCCTGCAGACAGCGGCAGCGAGTGAACATCGTGACCGACCACCAGT   4888

Qy    2174  ACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGGCGAGAGGCGAGCTGCTGGTGAACCAGA 2233
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    4889  ACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGGCGAGAGGCGAGCTGCTGGTGAACCAGA 4948

Qy    2234  TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 2293
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4949  TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 5008

Qy    2294  GCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCGCAGGGCATCTCCGGAGCAAGTGTGT 2353
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5009  GCATCGGCGGCAACGAGCAGATCGACAAGCTGGTGAGCGCAGGGCATCTCCGGAGCAAGTGTGT 5068

Qy    2354  TCCTGACGGCATCGATGGCGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG  2413
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5069  TCCTGACGGCATCGATGGCGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG 5128

Qy    2414  GCAGCGGCGGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGT   2463
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db    5129  GCAGCGGCGGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGT  5178
```

```
RESULT 12
US-10-190-435-13
;  Sequence 13, Application US/10190435
;  Publication No. US20030143248A1
;  GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT: BARNETT, Susan W.
;  APPLICANT: LIAN, Ying
;  APPLICANT: ENGELBRECHT, Susan
;  APPLICANT: VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE: 2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
;  SEQ ID NO 13
;  LENGTH: 3531
```

RESULT 13
US-10-190-435-14
; Sequence 14, Application US/10190435
; Publication No. US20030143246A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 3537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagPolmutAtt_C
US-10-190-435-14

Query Match   95.6%; Score 2361.2; DB 14; Length 3537;
Best Local Similarity   98.4%; Pred. No. 0;
Matches 2410; Conservative 0; Mismatches 28; Indels 12; Gaps 2;

```
Qy    254  AGGCCCGCGAGTTCCCAGGCAGCAGAAACCGGCCAACAGCCCGCCAACAGCCCACCAGCGCCGAGCTGC  313
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1334  AGGCCCGCGAGTTCCCCAGGCAGCGAGCAACCCGCCAACAGCAGCCCACCAGCGCCGAGCTGC 1393

Qy    314  AGGTCGCGGCGACAACCCCGCGAGCGCGAGGCGCGGCGAGCCCGAGCGCCAGGGCACCCTGAACT  373
Db   1394  AGGTGCGGCGGCACACCCCGCGAGCGCGAGGCGCGGCGAGCCCGAGCGCCAGGGCACCCTGAACT 1453

Qy    374  TCCCCAGATCACCCCTGTGGCAGCGGCGCGCCCCTGGTGCAGCAATCAAGGTGGGCGGCCAGATCA  433
Db   1454  TCCCCAGATCACCCCTGTGGCAGCGGCGCGCCCCTGGTGCAGCGATCATCAAGGTGGGCGGCCAGATCA 1513

Qy    434  AGGAGGCCCTGCTGGAACACCGGCCGGCCGCCGACGACACCGTGCTGGAGGAGGATGAGCTGCCCG  493
Db   1514  AGGAGGCCCTGCTGGACTCCGGCCGGCCGCCGACGACACCGTGCTGGAGGAGGATGAGCTGCCCG 1573

Qy    494  GCAAGTGGAAGCCCAAGATGATCGGCGGCCATCGGCCGCTTCATCGAAGGTGCGGCCAGTACG  553
Db   1574  GCAAGTGGAAGCCCAAGATGATCGGCGGCCATCGGCCGCTTCATCGAAGGTGCGGCCAGTACG 1633

Qy    554  ACCAGATCCTGATCGAGATCTCGCGGCAAGAGGCCATCGGCACCGTGCTGATCGGCCCCA  613
Db   1634  ACCAGATCCTGATCGAGATCTCGCGGCAAGAGGCCATCGGCACCGTGCTGATCGGCCCCA 1693

Qy    614  CCCCCTGAACATCGGCCGCCAGCTGGCTGCACCCTGAACTTCC  673
Db   1694  CCCCCGTGAACATCGGCCGCCAGCTGGCTGCACCCTGAACTTCC 1753

Qy    674  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGGCCGCATGGACGGCCCCAAGG  733
Db   1754  CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGGCCGCATGGACGGCCCCAAGG 1813

Qy    734  TGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCCCATCTGCGAGGAGA  793
Db   1814  TGAAGCAGTGGCCCCTGACCGAGGAAGATCAAGGCCCTGACCGCCCATCTGCGAGGAGA 1873

Qy    794  TGGAGAAGGAGGGCAAGATCACCAAAGATCGGCGAGAACAACCCCTACAACACCCCGTGT  853
Db   1874  TGGAGAAGGAGGGCAAGATCACCAAAGATCGGCGAGAACAACCCCTACAACACCCCGTGT 1933

Qy    854  TCGCCATCAAGAAGAAGACAGCACCAGTTCCGGCGCAAGCTGTGTGACTTCCGGGACTGA  913
Db   1934  TCGCCATCAAGAAGAAGACAGCACCAGTTCCGGCGCAAGCTGTGTGACTTCCGGGACTGA 1993

Qy    914  ACAAGCCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCACCCCGGCCGGCCTGA  973
Db   1994  ACAAGCCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCCCCACCCCGGCCGGCCTGA 2053

Qy    974  AGAAGAAGAAGAGAGCGGTGACCGTCGTGACCGTCGTCGACGTCGTGACCTGGCCCCCTGG 1033
Db   2054  AGAAGAAGAAGAGAGCGGTGACCGTCGTGACCGTCGTCGACGTCGTGACCTGGCCCCCTGG 2113

Qy   1034  ACGAGGACTTCCCGCAAGTACACCGGTGGGCAGCATCAACAACGAACACCCCGCCA 1093
Db   2114  ACGAGGACTTCCCGCAAGTACACCGGTGGGCAGCATCAACAACGAACACCCCGCCA 2173

Qy   1094  GCATCCGTCACCAGTACAAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 1153
Db   2174  GCATCCGTCACCAGTACAAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 2233

Qy   1154  AGAGACCATGACCAAGATCCTGGACGTGGACGTGCTGGAGCGTGGACGTGGAGCGTGATCT 1213
Db   2234  AGAGACAGCATGACCAAGATCCTGGACGTGGACGTGCTGGAGCGTGGACGTGGAGCGTGATCT 2293

Qy   1214  ACCAGTACATGGACGACCTGTACGTGGGCAGCCACCACCACGGCCGCCGCCA 1273
Db   2294  ACCAGTACATGGACGACCTGTACGTGGGCAGCAGCCACCACCACGGCCGCCGCCA 2347

Qy   2294  ACCA------GGCCCCGTAGTGGGCAGCAGCCACCACCACGGCCGCCGCCA 2347

Qy   1274  AGATCGAGGAGCTGCGCAAGCACCTGCTGCGGCTTCACCACCCCGACAAGAAGC 1333
Db   2348  AGATCGAGGAGCTGCGCAAGCACCTGCTGCGGCTTCACCACCCCGACAAGAAGC 2407
```

```
Qy   1334  ACCAGGAAGAGCCCCCTTCGTGGATGGGTACGAGCTGCACCCGACAAGTGGACCG 1393
Db   2408  ACCAGGAAGAGCCCCCTTCGTGCCCAT------CGAGCTGCACCCGACAAGTGGACCG 2461

Qy   1394  TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGAGTGGACCGTGAACGACATCCAGAAGCTGG 1453
Db   2462  TGCAGCCCCATCGAGCTGCCCGAGAAGGAGAGAGTGGACCGTGACGTGAACGACGTGG 2521

Qy   1454  TGGGCAAGCTGAACTGGGCGCCAAGGCCCTGCGCCGCATCAAGAGGTGCGGCCAGCTGTGCA 1513
Db   2522  TGGGCAAGCTGAACTGGGCGCCAGCCCAGGCCCTGCGCCGCATCAAGAGGTGCGGCCAGCTGTGCA 2581

Qy   1514  AGTCTGCGGCGGCGGCCAAGGCCCTGACCGACATGTGCCCCTGACCGGAGGAGCGGCCGAGC 1573
Db   2582  AGTCTGCTGGCGGCGGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGGAGGAGCGGCCGAGC 2641

Qy   1574  TGGAGCTGCCCGAGAACCTGCGCCGAGATCCTGCGCGAGCCCGTGCACGCGGTGTACTACGACC 1633
Db   2642  TGGAGCTGCCCGAGAACCTGCGCCGAGATCCTGCGCGAGCCCGTGACGCGGTGTACTACGACC 2701

Qy   1634  CCAGCAAGAACCTCGTGCCGAGATCAGCAGGCACAGCAGGGCCACCAGTTGGACCTACCAGA 1693
Db   2702  CCAGCAAGAACCTGGTGCCGAGATCAGCAGGCACAGCAGGGCCACCAGTTGGACCTACCAGA 2761

Qy   1694  TCTACCAGGAGCCCTTCAAGAACCTGAAGAACCTGGAGATGCGCCACCCGCACCCCCC 1753
Db   2762  TCTACCAGGAGCCCTTCAAGAACCTGAAGAACCTGGAGAGTGCGCAAGTGCGCCACCCGCCC 2821

Qy   1754  ACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGTGCAGAGATCGCCATGAGAGCATCG 1813
Db   2822  ACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGTGCAGAGATCGCCATGAGAGCATCG 2881

Qy   1814  TGATCTGGGCAAGACCCCCAAGTTCCGGCTGCCCATCCCGAGTGGGAGACCTGGGAGACCT 1873
Db   2882  TGATCTGGGCAAGACCCCCAAGTTCCGGCTGCCCATCCCGAGTGGGAGACCTGGGAGACCT 2941

Qy   1874  GGTGACCGACTACTGGCAGGCCCACCTGATCCCGAGTGGGAGTTCGTGAACACCCCCC 1933
Db   2942  GGTGACCGACTACTGGCAGGCCCACCTGATCCCGAGTGGGAGTTCGTGAACACCCCCC 3001

Qy   1934  CCCTGGTGAGCTGGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 1993
Db   3002  CCCTGGTGAGCTGGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCT 3061

Qy   1994  ACGTGGACGGCCGCGGCCAACCCGGCGAGACCAAGATCGGCAAGGCCGTACGTGACCGACC 2053
Db   3062  ACGTGGACGGCCGCGGCCAACCCGGCGAGACCAAGATCGGCAAGGCCGTACGTGACCGACC 3121

Qy   2054  GGGGCCGGCGCAGAAGATCGTTGAGCCTGACCGAGAGAAGACCGAGCCGAGCTGCGAGG 2113
Db   3122  GGGGCCGGCGCAGAAGATCGTTGAGCCTGACCGAGAGAAGACCGAGCCGAGCTGCGAGG 3181

Qy   2114  CCATCGGCGGCAACAGCTGGCCCTGCAGGACAGCAGCCAGCGCGAGGTGAACATCGTACAGCAGCCAGT 2173
Db   3182  CCATCGGCGGCAACAGCTGGCCCTGCAGGACAGCAGCCAGCGCGAGGTGAACATCGTGACCAGCAGCCAGT 3241

Qy   2174  ACGGCCCTGGGCATCCAGGCCCAGGCTGAAAGGCCAGCCCCAGGCAGTGGTGAACCAGA 2233
Db   3242  ACGGCCCTGGGCATCCAGGCCCAGGCTGAAAGGCCAGCCCCAGGCAGTGGTGAACCAGA 3301

Qy   2234  TCATCGAGCAGCTGATCAAGAAGCAGATCGTGATCT 2293
Db   3302  TCATCGAGCAGCTGATCAAGAAGCAGATCGTGATCT 3361

Qy   2294  GCATCGGCGGCAACGAGCAGATCGAGATCGGCACCGGGCCA 2353
Db   3362  GCATCGGCGGCAACGAGCAGATCGAGATCGGCACCGGGCCA 3421

Qy   2354  TCCTGGACGGGCATGGCGGCATCGTGATCTACGTGGGGCTGTG 2413
Db   3422  TCCTGGACGGGCATGGCGGCATCGTGATCTACGTGGGGCTGTG 3481

Qy   2414  GCAGCGGCGGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCCGGT 2463
```

```
Db  3482 GCAGCGGCGGCCCCTAGAGATCGATTAAAAGCTTCCCGGGGGCTAGCACCGGT 3531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 14
US-10-190-435-15
; Sequence 15, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT:  ZUR MEGEDE, Jan
; APPLICANT:  BARNETT, Susan W.
; APPLICANT:  LIAN, Ying
; APPLICANT:  ENGELBRECHT, Susan
; APPLICANT:  VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 3537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GagPolmutIna_C
US-10-190-435-15

Query Match           95.6%;  Score 2361.2;  DB 14;  Length 3537;
Best Local Similarity 98.4%;  Pred. No. 0;
Matches  2410;  Conservative     0;  Mismatches    28;  Indels    12;  Gaps     2;

Qy    14 TGGCCGAGGCCATGAGCCAGGCCACCGAGCCCAACATCCTGATGCAGCGCCAGCAACTCA 73
         ||||||||||| |||||||||||||| ||||||||||||||||||||||||||||| |||
Db  1094 TGGCCGAGGCGATGAGCGATGAGCCAGGCCAACACCGAGCCCAACATCCTGATGATGCAGCGCCAGCAACTTTAAAA 1153

Qy    74 AGGGCCCCAAGGAGCCATCATCGGCTTCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 133
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db  1154 AGGGCCCCAAGGAGCCATCATCAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 1213

Qy   134 ACTGCGCCGGCCCCCGCGCCCCCGAAGAGGGCCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1214 ACTGCGCCGGCCCCCGCGCCCCCGAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCGGCCACCAGATGA 1273

Qy   194 AGGACTGCACCGAGCGGCCAGCGCCAACTTCTTCCGCGAGGAACCGGCCAACAGCCGCCCCAGCCCAGGGCA 253
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1274 AGGACTGCACCGAGCGGCGCAGCGCCAACTTCTTCCGCGAGGAACCGGCCAACAGCCGCCCCGAGCCCAGCCCAGGGCA 1333

Qy   254 AGCCCGGCGAGTTCCCCCAGGCACAACCCCCGACGAGCCGGAGCCGGAGCTGC 313
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1334 AGCCCGGCGAGTTCCCCCAGGCAGACAACCCCGACGAGCCGGAGCCGGAGCTGC 1393

Qy   314 AGGTGCGCGGCGGCACACCCCTGTGCAGCGCCCCCTGTGACTCAAGGTGGGCGGCCAGATCA 373
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1394 AGGTGCGCGGCGGCGACAACCCCCTGTGCAGCGCCCCCGAGCCGGCCCCTGTGACTCAAGGTGGGCGGCCAGATCA 1453

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGACCATCAAGGTGCTGCGGCGGCCAGATCA 433
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1454 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGACCATCAAGGTGCTGCGGCGGCCAGATCA 1513

Qy   434 AGGAGGCCCTGCTGGACACCCGGCGGCCGCGACGCGCCGAGCGGGGAGATGAGCCTGCCCG 493
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1514 AGGAGGCCCTGCTGGACACCGGCGGCCGCGACGCGCCGAGCGGGGAGATGAGCCTGCCCGAGC 1573

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGGCAGTACG 553
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1574 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGGCAGTACG 1633

Qy   554 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGGTCGGCCCCA 613
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1634 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGGTCGGCCCCA 1693

Qy   614 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1694 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 1753

Qy   674 CCATCAGCCCCATCGAGACCGTGCCGGTGAAGCTGAAGCCGGGCATGGACGGCCCCAAGG 733
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1754 CCATCAGCCCCATCGAGACCGTGCCGGTGAAGCTGAAGCCGGGCATGGACGGCCCCAAGG 1813

Qy   734 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 793
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1814 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 1873

Qy   794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 853
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1874 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 1933

Qy   854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1934 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 1993

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 973
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1994 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 2053

Qy   974 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCGCTACTTCAGGGTGCCCCTGG 1033
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2054 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCGCTACTTCAGGGTGCCCCTGG 2113

Qy  1034 ACGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACACCGAGACCCCCG 1093
         ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db  2114 ACGAGACTTCCGCAAGTACACAGCCTTCACCATCCCCAGCATCAACACCGAGACCCCCG 2173

Qy  1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCC 1153
         |||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  2174 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 2233

Qy  1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 1213
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2234 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 2293

Qy  1214 ACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCA 1273
         ||||||                ||||||||||||||||||||||||||||||||||||||
Db  2294 ACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCA 2347

Qy  1274 AGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAAGC 1333
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2348 AGATCGAGGAGCTGCGCCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAAGAAGC 2407

Qy  1334 ACCAGAAGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCGACAAGTGACCG 1393
         |||||||||||||||||||||||||||||||       ||||||||||||||||||||
Db  2408 ACCAGAAGGAGCCCCCCTTCCTGTGCCAT-------CGAGCTGCACCCGACAAGTGACCG 2461

Qy  1394 TGGGCCCATGAGCTGCCGGAGGAGAGCTGAACGACATCCAGAGAGCTGG 1453
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2462 TGGGCCCATGAGCTGCCGGAGGAGAGCTGAACGACATCCAGAGAGCTGG 2521

Qy  1454 TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA 1513
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2522 TGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCA 2581

Qy  1514 AGCTGCTGCGGGGCACCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGC 1573
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2582 AGCTGCTGCGGGGCACCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGC 2641

Qy  1574 TGGAGCTGGCCGAGAACCGGGAGATCCTGGGGCTTCATCAAGGTGCGGGTGTTACTACGACC 1633
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2642 TGGAGCTGGCCGAGAACCGGGAGATCCTGGGGCTTCATCAAGGTGCGGGTGTTACTACGACC 2701

Qy  1634 CCAGCAAGGACCTGTGGGGCTCGGCCAAGAAGCAGGCCACGAGCGGGCACTACCAGA 1693
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2702 CCAGCAAGGACCTGGTTGGCGGGCCCGAGATCGAGCAGGAGGCCACGAGCGGGACTACCAGA 2761

Qy  1694 TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGGCACCGCGCCCCCCA 1753
```

```
;  FEATURE:
;  OTHER INFORMATION: Description of Artificial Sequence:
;  OTHER INFORMATION: GagComplPolmutInaTatRevNef_C
US-10-190-435-12

Query Match       95.6%;    Score 2360.2;    DB 14;   Length 5145;
Best Local Similarity  99.2%;   Pred. No. 0;
Matches 2397;   Conservative   0;   Mismatches   8;   Indels  12;   Gaps  2;

Qy     14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACTTCA 73
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1487 TCGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAGCGCCAGCAACTTCA 1546

Qy     74 AGGGCCCCAAGCCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1547 AGGGCCCCAAGCCGCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCCGCA 1606

Qy    134 ACTGCCGGCCGCCCCCGCAAGAAGGGCTGTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 ACTGCCGGCCGCCCCCGCAAGAAGGGCTGTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 1666

Qy    194 AGGACTGCACGCCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA 253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1667 AGGACTGCACGCCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA 1726

Qy    254 AGGCCCGCGAGTTCCCAGCGAGACGAGAACCGCCGCCAGAGCCCCAGCCCGGCGGAGCTGC 313
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1727 AGGCCCGCGAGTTCCCAGCGAGACGAGAACCGCCGCCAGAGCCCCAGCCCGGCGGAGCTGC 1786

Qy    314 AGGTGCGCGGCGACACCCCCGCGCCGAGGAGGCCAGGCCGCCAGGGCCAGGCCACCCTGAACT 373
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1787 AGGTGCGCGGCGACACCCCCGCGCCGAGGAGGCCAGGCCGCCAGGGCCAGGCCACCCTGAACT 1846

Qy    374 TCCCCCAGATCACCCTGTGCGAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1847 TCCCCCAGATCACCCTGTGCGAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCA 1906

Qy    434 AGGAGGCCCTGCTGACCACCGGCGCGCCGCCGAGATGAGCGCTGCCCG 493
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1907 AGGAGGCCCTGCTGACCACCGGCGCGCCGCCGAGATGAGCGCTGCCCG 1966

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCGCTTCATCAAGGTGCGCCAGTACG 553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCGCTTCATCAAGGTGCGCCAGTACG 2026

Qy    554 ACCAGATCCTGATCGAGATCTGCGGGCACAAGGCCATCGGCACCGTGCTGTGATCGGCCCCCA 613
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2027 ACCAGATCCTGATCGAGATCTGCGGGCACAAGGCCATCGGCACCGTGCTGTGATCGGCCCCCA 2086

Qy    614 CCCCGTGACATCATCGGCCGCAACATGTCGCCGTGGGCTTGCACCTGGACTTCC 673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2087 CCCCGTGACATCATCGGCCGCAACATGTCGCCGTGGGCTTGCACCTGGACTTCC 2146

Qy    674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGAGCGGCCCCAAGG 733
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2147 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGAGCGGCCCCAAGG 2206

Qy    734 TGAAGCAGTGCCCCTGACCGAGGAAGAATCAAGGCCCTGACCGGCCATCTGCGAGGAGA 793
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2207 TGAAGCAGTGCCCCTGACCGAGGAAGAATCAAGGCCCTGACCGGCCATCTGCGAGGAGA 2266

Qy    794 TGGAGAAGGAGGGGCGAAGATCACCAAGATCGGCCCCGAGACCCCCTACAACACCCCCGTGT 853
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2267 TGGAGAAGGAGGGGCGAAGATCACCAAGATCGGCCCCGAGACCCCCTACAACACCCCCGTGT 2326

Qy    854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 2386

Qy    914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTCGGCATCCCCCACCCCGCCGGCCTGT 973
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTCGGCATCCCCCACCCCGCCGGCCTGA 2446

Db   2762 ...TCTACCAGGAGCCCTTCAAGAACCTGACCGAACGCGCAAGTACGCCGCCAAGATGCGCCAGTACGCCGCCATCG 2821
Qy   1754 ACACCAACGACGTGAAGCAGCAGTGCAGCAGC... 1813
Db   2822 ACACCAACGACGTGAAGCAGCAGTGCAGCAGC... 2881
Qy   1814 TGATCTGGTGGCAAGACCCCCAAGTTCCGCCTGCCCATCCGAAGGAGGAGACCTGGGAGACCT 1873
Db   2882 TGATCTGGTGGCAAGACCCCCAAGTTCCGCCTGCCCATCCGAAGGAGGAGACCTGGGAGACCT 2941
Qy   1874 GGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCC 1933
Db   2942 GGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCC 3001
Qy   1934 CCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGAGAGCCCATCATCGGCGCCGAGACCTTCT 1993
Db   3002 CCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGAGAGCCCATCATCGGCGCCGAGACCTTCT 3061
Qy   1994 ACGTTGGACGGGCGCGCCCCAACCGCCAACCGCGAGATCGGCAAGGCCGGCTACGTGACCGACC 2053
Db   3062 ACGTTGGACGGGCGCGCCCCAACCGCCAACCGCGAGATCGGCAAGGCCGGCTACGTGACCGACC 3121
Qy   2054 GGGGCCGGCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGCCCGGCGGAGCTGCAG 2113
Db   3122 GGGGCCGGCGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGCCCGGCGGAGCTGCAG 3181
Qy   2114 CCATCCAGTGCGCCCTGCCAGGACAGCGGCGCAGCGGCGAGGTGAACATCGTGACCGACACAGCCAGT 2173
Db   3182 CCATCCAGTGCGCCCTGCCAGGACAGCGGCGCAGCGGCGAGGTGAACATCGTGACCGACACAGCCAGT 3241
Qy   2174 ACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCCCGAGCAGAGCGAGACGAGCTGGTGAACCAGA 2233
Db   3242 ACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAGCCCGAGCAGAGCGAGACGAGCTGGTGAACCAGA 3301
Qy   2234 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 2293
Db   3302 TCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 3361
Qy   2294 GCATCGGGGCCAACGAGCAGATCGACAACGAGCAGATCGGCATCCGCAAGGTGCTGT 2353
Db   3362 GCATCGGGGCCAACGAGCAGATCGACAACGAGCAGATCGGCATCCGCAAGGTGCTGT 3421
Qy   2354 TCCTGGACGGCATCGATGGCGACGACTGTGATCGTCGTGACACTGTACCAGTACGTGG 2413
Db   3422 TCCTGGACGGCATCGATGGCGACGACTGTGATCGTCGTGACACTGTACCAGTACGTGG 3481
Qy   2414 GCAGCGGCGGCGGCCCCTAGGATCGGATTAAAAAGCTTCCCGGGGCTAGCACCGGT 2463
Db   3482 GCAGCGGCGGCGGCCCCTAGGATCGGATTAAAAAGCTTCCCGGGGCTAGCACCGGT 3531
```

RESULT 15
US-10-190-435-12
; Sequence 12, Application US/10190435
; Publication No. US20030143248A1
GENERAL INFORMATION:
   APPLICANT: ZUR MEGEDE, Jan
   APPLICANT: BARNETT, Susan W.
   APPLICANT: LIAN, Ying
   APPLICANT: ENGELBRECHT, Susan
   APPLICANT: VAN RENSBURG, Estelita J.
   TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
   TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
   FILE REFERENCE: PP.8133.003 / 2302-18133
   CURRENT APPLICATION NUMBER: US/10/190,435
   CURRENT FILING DATE: 2002-12-30
   NUMBER OF SEQ ID NOS: 319
   SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
; LENGTH: 5145
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
Qy   974 AGAAGAAGAGAGCGGTGACCGTGCTGGCGAGCGCCTACTTCAGCGTGCCCCTGG 1033
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2447 AGAAGAAGAGAGCGGTGACCGTGCTGGCGAGCGCCTACTTCAGCGTGCCCCTGG 2506

Qy  1034 ACCAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACACGAGCCCCCG 1093
             ||||||||||||||||||||| |||||||||||||||||||||||||||||| |||
Db  2507 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACACGAGCACCCCG 2566

Qy  1094 GCATCCGGTCACCAGTACAAACGTGCTGCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC 1153
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2567 GCATCCGGTCACCAGTACAAACGTGCTGCCCAGGGCTGGAAGGGCAGCCCAGCATCTTCC 2626

Qy  1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCCAACCCCGAGATCGTGATCT 1213
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2627 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGGCCCAACCCCGAGATCGTGATCT 2686

Qy  1214 ACCAGTACATGGACGACCTGTACGTGGGCAGCGACCTGGCAGCGGCACCGCGCCA 1273
             |||||||||                    |||||||||||||||||||||||||
Db  2687 ACCA------GGCCCCCCCGTACGTGGGCAGCGACCTGGCAGCGGCACCGCGCCA 2740

Qy  1274 AGATCGAGGAGCTGCGCAAGCACCTGTCGCGTGGGGCTTCACCACCCCGACCAAGAAGC 1333
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2741 AGATCGAGGAGCTGCGCAAGCACCTGTCGCGTGGGGCTTCACCACCCCGACCAAGAAGC 2800

Qy  1334 ACCAGAAGAGACCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCG 1393
             |||||||||||||||||||||||||||||              |||||||||||||||
Db  2801 ACCAGAAGAGACCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACAAGTGGACCG 2854

Qy  1394 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGCATCCAGAAGCTGG 1453
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2855 TGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGCATCCAGAAGCTGG 2914

Qy  1454 TGGGCAAGCTGAACTGGGCGGCCAAGGCCCTGAGATCTACCCCGGCCATCGTGCCCAGCTGTGCA 1513
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2915 TGGGCAAGCTGAACTGGGCGGCCAAGGCCCTGAGATCTACCCCGGCCATCGTGCCCAGCTGTGCA 2974

Qy  1514 AGCTGCTGCGCGGCGGCCAAGGCCCGAGATCCTGCGCGAGCCCGTGCGCAGCCGAGGAGGCCGAGC 1573
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2975 AGCTGCTGCGCGGCGGCCAAGGCCCGAGATCCTGCGCGAGCCCGTGCGCAGCCGAGGAGGCCGAGC 3034

Qy  1574 TGGAGCTGCCGGAGAACCGCAGCTGACCGAGATCCTGCGCCGAGCCCGTGCACGGCCGTGTACTACGACC 1633
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3035 TGGAGCTGCCGGAGAACCGCAGCTGACCGAGATCCTGCGCCGAGCCCGTGCACGGCCGTGTACTACGACC 3094

Qy  1634 CCAGCAAGGACCTGTGCGGCGAGATCGTGTGGGCCAGAAGCAGGGGCCACAGGACCAGTGGACCTACCAGA 1693
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3095 CCAGCAAGGACCTGTGCGGCGAGATCGTGTGGGCCAGAAGCAGGGGCCACAGGACCAGTGGACCTACCAGA 3154

Qy  1694 TCTACCAGGAGCGCCCTTCAAGAACCTGAGAACCGGCAAGTACGCCCAGCAGTGCCAAGATGCCACCGCCC 1753
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3155 TCTACCAGGAGCCCTTCAAGAACCTGAGAACCGGCAAGTACGCCCAGCAGTGCCAAGATGCCACCGCCC 3214

Qy  1754 ACACCAACGACGTGAAGCAGCTGAACCGCGACCGACCTGCAGAAGATCGGCCATGGAGAGCATCG 1813
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3215 ACACCAACGACGTGAAGCAGCTGAACCGCGACCGACCTGCAGAAGATCGGCCATGGAGAGCATCG 3274

Qy  1814 TGATCTGGGCAAGACCCCCCAAGTTCCGCTGCCCAAGTTCCGCTGCCCTGCCCATCCAGAGGAGACCTGGGAGACCT 1873
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3275 TGATCTGGGCAAGACCCCCCAAGTTCCGCTGCCCAAGTTCCGCTGCCCTGCCCATCCAGAGGAGACCTGGGAGACCT 3334

Qy  1874 GGTGGACCACTACTGGCAGGCCACCTGGCAGGCCCACCTGCGAGTTGGGAGTTCGTGAACACCCCCC 1933
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3335 GGTGGACCACTACTGGCAGGCCACCTGGCAGGCCCACCTGCGAGTTGGGAGTTCGTGAACACCCCCC 3394

Qy  1934 CCCTGGTGAAGCTGGTGGTACCAGCTTGGAGAAGGAGAAGATGGCCCATCATCGGCGGCCCGAGACCTTCT 1993
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3395 CCCTGGTGAAGCTGGTGGTACCAGCTTGGAGAAGGAGAAGATGGCCCATCATCGGCGGCCCGAGACCTTCT 3454

Qy  1994 ACGTTGGACGGCGCGGCCGAGATCGGCCAAGAGGCCGGCTTACGTGACCGACC 2053
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3455 ACGTTGGACGGCGCGGCCGAGATCGGCCAAGAGGCCGGCTTACGTGACCGACC 3514

Qy  2054 GGGGCCGGCCGGAAGAGATCGTGAGCCTGACCGGACCACCGAGAAGACCAGCCGAGCTGCAGG 2113
```

```
Db  3515 GGGGCCGGCCGGAAGAGATCGTGAGCCTGACCGGACCACCGAGAAGACCGAGCTGCAGG 3574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  2114 CCATCCAGCTGGCCCTGCAGGACAACGGCGGCGAGGTGAACATCGTGACCGACGCCAGT 2173
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3575 CCATCCAGCTGGCCCTGCAGGACAGCGGCGGCGAGGTGAACATCGTGACCGACGCCAGT 3634

Qy  2174 ACGCCCTGGGCATCATCCAGCCCGACAAGAGCGAGAGCTGGTGAACCAGA 2233
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3635 ACGCCCTGGGCATCATCCAGCCCGACAAGAGCGAGAGCTGGTGAACCAGA 3694

Qy  2234 TCATCGAGCAGCTGATCAAGAAGGAGAAGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 2293
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3695 TCATCGAGCAGCTGATCAAGAAGGAGAAGTGTACCTGAGCTGGGTGCCCGCCCACAAGG 3754

Qy  2294 GCATCCGGCGGCAACAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGT 2353
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3755 GCATCCGGCGGCAACAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGT 3814

Qy  2354 TCCTGGACGGCATCGATGGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG 2413
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3815 TCCTGGACGGCATCGATGGGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGG 3874

Qy  2414 GCAGCGCGGCGCCCTAGG 2430
             ||||||||||||||||||
Db  3875 GCAGCGCGGCGCCCTAGG 3891
```

Search completed: April 10, 2004, 21:24:32
Job time : 604.779 secs

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       April 10, 2004, 06:34:42 ; Search time 4127.92 Seconds
                                         (without alignments)
                                         17774.420 Million cell updates/sec

Title:        US-09-610-313-32
Perfect score: 2457
Sequence:     1 gtcgacgccaccatggccga.........gggctagcaccggtgaattc 2457

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:     27513289 seqs, 1493109276 residues

Total number of hits satisfying chosen parameters:    55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    EST:*
          1:  em_estba:*
          2:  em_esthum:*
          3:  em_estin:*
          4:  em_estmu:*
          5:  em_estov:*
          6:  em_estpl:*
          7:  em_estro:*
          8:  em_htc:*
          9:  gb_est1:*
         10:  gb_est2:*
         11:  gb_htc:*
         12:  gb_est3:*
         13:  gb_est4:*
         14:  gb_est5:*
         15:  em_estfun:*
         16:  em_estom:*
         17:  em_gss_hum:*
         18:  em_gss_inv:*
         19:  em_gss_pln:*
         20:  em_gss_vrt:*
         21:  em_gss_fun:*
         22:  em_gss_man:*
         23:  em_gss_mus:*
         24:  em_gss_pro:*
         25:  em_gss_rod:*
         26:  em_gss_phg:*
         27:  em_gss_vrl:*
         28:  gb_gss1:*
         29:  gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	90.2	3.7	2598	11	AY103647	AY103647 Zea mays
2	85.6	3.5	951	12	BM321451	BM321451 rockefell
c 3	80.6	3.3	869	14	CK159167	CK159167 FGAS04056
c 4	80.2	3.3	1132	12	BM320864	BM320864 rockefell

5	79.6	3.2	1165	12	BM320900	BM320900 rockefell
6	79	3.2	867	12	BM321430	BM321430 rockefell
7	76.6	3.1	1550	12	BM321022	BM321022 rockefell
8	75	3.1	545	12	BI724851	BI724851 1031075E0
9	75	3.1	862	12	BM321023	BM321023 rockefell
10	74.4	3.0	914	28	BZ568300	BZ568300 pacs2-164
11	73.8	3.0	2433	29	AY401196	AY401196 Homo sapi
12	73.2	3.0	853	12	BM321393	BM321393 rockefell
c 13	72.8	3.0	566	14	BM587428	BM587428 170006873
14	72.6	3.0	788	14	CB641171	CB641171 OSJNEb03L
15	72.4	2.9	753	29	CC675888	CC675888 OGWCO51TH
16	72.2	2.9	640	14	BE601575	BE601575 HVSMEh009
17	72.2	2.9	688	14	CB648640	CB648640 OSJNEb12C
18	72.2	2.9	764	14	CB651670	CB651670 OSJNEb16L
19	72.2	2.9	766	14	CB642928	CB642928 OSJNEb03F
20	72.2	2.9	809	14	CB641397	CB641397 OSJNEb01A
21	72.2	2.9	841	14	CB651502	CB651502 OSJNEb16H
c 22	72.2	2.9	851	29	CG260623	CG260623 OGWGE95TV
23	71.8	2.9	1304	28	BZ577729	BZ577729 msh2_5546
24	71.8	2.9	1933	11	AY325173	AY325173 Rattus no
25	71.6	2.9	500	12	BM372120	BM372120 EBro03_SQ
26	71.6	2.9	538	12	BM368580	BM368580 BBem08_SQ
27	71.6	2.9	540	9	AJ471121	AJ471121 AJ471121
28	71.6	2.9	566	13	BQ464692	BQ464692 HF02P20r
29	71.6	2.9	575	13	BU976068	BU976068 HA03B08r
30	71.6	2.9	576	13	BU984666	BU984666 HF04K03r
31	71.6	2.9	578	12	BM377112	BM377112 EBem05_SQ
32	71.6	2.9	579	10	BF253640	BF253640 HVSMEf000
33	71.6	2.9	582	12	BM372871	BM372871 EBma04_SQ
34	71.6	2.9	586	13	BU994273	BU994273 HM06H03r
35	71.6	2.9	588	13	BQ765498	BQ765498 EBro03_SQ
36	71.6	2.9	596	13	BU997745	BU997745 HI08O07r
37	71.6	2.9	600	9	AL508778	AL508778 AL508778
38	71.6	2.9	615	13	CA009718	CA009718 HU14O18r
39	71.6	2.9	621	13	BU986998	BU986998 HF13108r
40	71.6	2.9	625	13	BQ768295	BQ768295 EBro08_SQ
41	71.6	2.9	637	13	BU999333	BU999333 HI14D01r
42	71.6	2.9	641	13	BU978712	BU978712 HA14B02r
43	71.6	2.9	645	13	BU977527	BU977527 HA11U06r
44	71.6	2.9	646	13	BU968956	BU968956 HB10B12r
45	71.6	2.9	655	13	BU996444	BU996444 HM13K04r

ALIGNMENTS

```
RESULT 1
AY103647
LOCUS       AY103647                2598 bp    mRNA      linear    HTC 16-OCT-2002
DEFINITION  Zea mays PCO142084 mRNA sequence.
ACCESSION   AY103647
VERSION     AY103647.1  GI:21206725
KEYWORDS    HTC.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 2598)
  AUTHORS   Hainey,C.F., Dolan,M., Miao.G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2  (bases 1 to 2598)
  AUTHORS   Coe,E.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
```

```
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES             Location/Qualifiers
     source          1..2598
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:638378"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Cornsensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
ORIGIN

Query Match      3.7%;   Score 90.2;  DB 11;  Length 2598;
Best Local Similarity   42.2%;  Pred. No. 0.24;
Matches  814;  Conservative   0;  Mismatches 1098;  Indels   15;  Gaps    5;
```

```
Qy    296  CCACCACCGCCGGAGCTGCAGTGCGCGGCGACAACCCCCGCAGCAGGGCCGGCGGCGAGC 355
Db     51  CCACCACCTCCACCTTCCTGCCACCGCGGCGCCCCAACCAACCACACGGAGCGAGGCGCAA 110

Qy    356  GCCAGGGCACCCGAACTTCCCCAGATCACCCGTGTGGCAGCGCCCCTTGTGGAGCATCA 415
Db    111  TGGCGGGCAACGGCGCCATCGTGGAGGAGCGACCCCGCTGAACTGGGGCGCGGCGGCGGCGG 170

Qy    416  AGGTGGCGGCCAGATCAAGGAGGCCCCTGCTGGACACCGACCACCGTGCTGG 475
Db    171  AGCTGGCCGGGAGCCACCTGACGCAGGTGAAGCGAGGTGAAGCGATGGTGGCGCAGCCCG 230

Qy    476  AGGAGATGAGCCTGCCCGGCAAGTGAAGCCGGCCAAGATGATCGGCGGCATCGGCGGCTTCA 535
Db    231  TGGTCAAGATCGAGGGTTCCACCCCTCGGCCAGGTGGCCGGTCGGCGTCGCCTTCCGCCA 290

Qy    536  TCAAGGTGCGCCAGTACGACCCAGATCCTGATCGAGATCTGCGCGCAAGAAGGCCATCGGCA 595
Db    291  AGGACGCGGTCCGGCGTCCGGAGCTCGAGCGTCGAGACGGAGGGCCCGCCGGGCGTCAAGGCCA 350

Qy    596  CGGTGCTGATCGGCCCCGTGAACATCATCGCCGGTGTGGACCAGCTGG 655
Db    351  GCAGCGAGTGGATCCTCGAGTCGATGCTGCCCCACGCCGCCGGACCATCTTACGGCGTCACCACCG 410

Qy    656  GCTGCACCGTGAACTTCCCCATCAGCCCGTGCCGGCCCGACGGAGGAGGAGATCAAGGCCCTGA 715
Db    411  GCTTCGGCGGCACCTGCCACCCTCCACCGGCGACGGGACCCGCTGCCGGTCCCAGGTCGGCTGC 470

Qy    716  GCATGACGGCGCCCAAGGTGAAGCACGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGA 775
Db    471  TCAGGCAGCATCTCAACGCGCGGAATCTTCGGCACGAGGGCAAGATCACCAAAGATCGGCGG 530

Qy    776  CCGCCATCATCTGCGAGGAGAGATGGAGGAGGGCCATCATCGCCGGATGCGGGTG 835
Db    531  AGGGTCACCCGCGCGGGATGCTGGTGTCGCGCATCACGAGGCTACTCGGCG 590

Qy    836  CCTACAACACCCCGTGTTCCCGCCATCAAGAAGGACAGCACCAAGATGGCGCAAGCTGG 895
Db    591  TCCGCTTCGAGATCCTGAGGCCATCACGAGCTGCTCAACACCGGTGTCAGCCCCTGCC 650

Qy    896  TGGACTTCCGGCAGCTGAACACAGCTGAAACGGCCACCCAGGACTTCTGGGAGGTGAGCAGCTGGGCATCC 955
Db    651  TGCCGCTTCGGGCACGCCACCATCACCGGCGTCGGGGCGAACCTGGTCCCGCTCTCCTACATCGCCG 710

Qy    956  CCCACCCCGCCGGCCTGAGAGAAGAAGAAGAGAGAGCGTGACCTGCTGGACGTGGCGCACGCCT 1015
Db    711  GCCTCATCATCACGGGCGCCCCACCCGCCGCCGGCCGTCACCG---TCGACGCGGAGGAGGTGG 767

Qy   1016  ACTTCAGCGTGCCCCTGGACCGAGGACTTCCGCAAGTACACCGCCCTTCACCGCCGGCACA 1075
Db    768  ACGCCGCGAGGGGCGTTCAAGATCGCCTCAAGATCGCCCGGGCATCGGGGGCCGGCTTCTCTTCAAGGCTCAACCCCA 827

Qy   1076  TCAACAACGAGACCCCGCCATCCGTCACCGGCACGTCCGTGGGCTCCGCGCTCGCGGCCACCGTGGAAGG 1135
Db    828  AGGAGGGCCTCGCCGCATCCCGCCATCGTCAACGGCACGTCCGGCGTCCGCGCTCGCGGCCACCGTGA 887

Qy   1136  GCAGCCCCAGCATCTTCAGAGCAGCAGCAAGATCGACAAGATCCTGGAGCCCTTCCGCGCCCGCA 1195
Db    888  TGTACGAGCCAACGTCCTGGCCGTCGGCGAGGTCCT---GTCCCGCGTCTTCGTCG 944

Qy   1196  ACCCCGAGATCGTGATCTACCAGGCCCCCGTGTACGTGGGCAGCGACCTGGAGATCGGCC 1255
Db    945  AGGTCATGAACGGCAAGCCCGACTACACGACCACCTGACCCAAGCTGAAGCACCACC 1004

Qy   1256  AGCACCGCCGCCAGATCGAGGAGCTGCCCAAGCTGCACCACCTGCTG---CGTTGGGCGTTCACCA 1312
Db   1005  CGGGGTCCATCGAGGCGGCCCATCATGAGGCACACTCCTGGATGGCAGCCTCCTTCATGA 1064

Qy   1313  CCCCCGACAAGAGAGCACCCAGAAGGAGCACCCCCTTCTGCGCCATCGAGCTGCACCCCGACA 1372
Db   1065  AGCAGGCCAAGAAGTTGAACGAGCTGGACCCGGTTGGAAGCCCAAGCAGGAGCAGGTACG 1124

Qy   1373  AGTGGACCGTGCGAGCCCATCGAGCTGCCCGAGATCGAAGAGGAGAGCTGGACCTGAACGACATCC 1432
Db   1125  CGCTCCGCACGTCCGCAGTGGCTGGCGGCAGCCAGATCCATCGGCGCGGCGCCACCA 1184

Qy   1433  AGAAGCTGGTGGGCAAGCTGAACTGGCTG---GCTGGGCAGCGCCAGATCTCACCCGGCATCAAGGTGCGCC 1492
Db   1185  AGTCCATCGAGCGCGAGGTCAACTCCGTGAACGA---CAACCCGGTCATCGACGTCCACC 1241

Qy   1493  AGCTGTGCAAGCTGCTGCGCGGCGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGG 1552
Db   1242  GCGGGCAAGGGCGTTGCACGGCGGCCAACTTCTCAGGCACCCCATCGGGCGTGTCCATGGACA 1301

Qy   1553  AGGCCGAGCTGGAGCTGCCCGAGAACCGGCCAAGATCGGCGCGGCCGTGTGCGCGCGGCTTCA 1612
Db   1302  ACGCCCGCCTCGGCCATCGGCAAGCTCATGTTCCGCGAGCTCGAGCTCG 1361

Qy   1613  ACTACGACCCCGAAGGACCCCGTGCTGGCCGAAGAGAGGCCATCGGCA 1672
Db   1362  TCAACGAGTTCTACAACAACGGGCTCACCTCCAACCTGGCGCCAGCCCCAGCC 1421

Qy   1673  CCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGAACCGGCCAAGTACGCCAAGATGC 1732
Db   1422  TGGACTACGGCTTCAAGGACCCATGGCCGACCGCGTGCGAGCTCTCGAGCTCCAGT 1481

Qy   1733  GCACCGCCCACCAACCAACCAGTGAAGAGCAGCTGACCAGCCGTGCAGAAGTGCCATGG 1792
Db   1482  ACCTGGGCAACCCCATCACCAACAACCAGTGACAGCGCGCGAGCAGCACAACCAGGACGTGA 1541

Qy   1793  AGAGCAT---CGTGATCTGGGGCAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGAGAGA 1849
Db   1542  ACTCCCTGGGCCTGTCTCGGCCAGGAAGACCGGCCAGGGCGATCGAACATCCTGAAGCTCA 1601

Qy   1850  CCTGGGAGACCTGGGCGGACCACCGAGCTACTGGGACGCGCGCCACCTGATCCCCGAGTGGGAGTTCG 1909
Db   1602  TGTGTCCACCTACATCGTGGCGCCTGTCCCAGCCGTGCCCAGCCGCGCCACCTGCGCGCACCTGA 1661

Qy   1910  TGAACACCCCCCCCGTGAAGCTGTGTGTACAGCTGGAGAGGAGGCCCATCATCGGGCG 1969
Db   1662  ACATCAAGGCGTCGGTGAAGAACACCGGTGCCCAGCCGGCCAAGAAGGTGCCGGCCATGA 1721

Qy   1970  CCGAGACCTTCTACGTGACCGGGCCGCGCCAACCGCCGAGACCAAGGCCGGCT 2029
Db   1722  ACCCCTCGGGCGAGCTTCCCAGCCGCGCCCGTTCAGCGAGGGACTCGATCAGCCCCATCG 1781

Qy   2030  ACGTGACCACCGGGGCCCGGGCCGCCCAGAAGATCGTGAGCGGAGGACCGACGACCAACCAGAGA 2089
Db   1782  ACCGGGAGGGCCGTGTTCACGTACGGCGGAGGACGGCGCAGCCGCGCAGCCGGCTGCCGGTGATGC 1841

Qy   2090  CCGAGCTGACGGTGCCCCTGGACCGAAGTACACCGCCCTTCACCGCCGCCAGCA 2149
Db   1842  AGAAGCTGGCGCGCCTGCGTGCTGGTGGACGCTACACCGCCCTGCACCATCCCCAGCA 1901
```

Qy   2150 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGACAAGAGCGAGAGCCGAGC 2209
           |||| || ||||||
Db   1902 CCTCCGTGTTCTCCAAGATCACCAGGTTCGAGGAGGAGTCCCGCGCGGTGCTGCCCCAGG 1961

Qy   2210 TGGGTGAA 2216
           ||||||
Db   1962 AGGGTGGA 1968

RESULT 2
BM321451
LOCUS   BM321451   951 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.1211 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION   BM321451
VERSION   BM321451.1 GI:18055857
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 951)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
   MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 951 Std Error: 0.00
            POLYA=No.
FEATURES   Location/Qualifiers
     source   1..951
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.5%; Score 85.0; DB 12; Length 951;
Best Local Similarity 46.0%; Pred. No. 0.69;
Matches 323; Conservative   0; Mismatches 376; Indels   3; Gaps   1;

Qy   1722 CGCCAAGATGCGCACCGCCACCGCCCCACCAACGACGTGAAGCAGTGACGAGGCCGTGCAGAA 1781
           || | | | || || |||||| || || ||||
Db   219 CGCCAAGATCCGCTGCGTCGTGCAACATCTTCTCGACGAGGAGACCACGCGGGCCGGCCGC 278

Qy   1782 GATCGCCATGAGAGCATCGTGATCTGGGCAAGACCCCCAAGTTCCGCGCCTGCCCATCCA 1841
              ||| || || ||| || ||| |||||| ||||| || |||
Db   279 CATCGGCGCAGGCGACGGCGTCTCGGTCTTCGCCTGGAAGGGCGGAGAACCTCCAGGAGTACTG 338

Qy   1842 GAAGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTG 1901
            || || || ||| ||| ||| |||| ||||| ||| |||||| ||||
Db   339 GGAGTGCACCTGGAAGGCCCTGTGCTTCGGCCTTCCCACCAGGGCCCCTCAGATCATCGTCGA 398

Qy   1902 GGAGTTCGTGAACAACACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGAGGCCCAT 1961
            |||||| || || ||||||| | | || || || ||| ||||||||||| | |
Db   399 CGACGGCGGTGACGCCACTCGATGATCCACAAGGGGTTCGCGGCCAGGACAACCCCCA 458

Qy   1962 CATCGGCGCGCCGAGACCTTCTACGCGTGACGGCGCCCCAACCCAGGCCGAGACCAAGATCGGCAA 2021
            |||| || ||||| | | || || || |||| || | ||||| |||
Db   459 GCTGCTGGAGGACGACGAGGGCCTCGAGGAGGTCGCCTCGCCTCCTCAACAACGTGCTCAAGCA 518

Qy   2022 GGCCGGCTACGTGACCGACCGACCGGCCGCGGGCCGGCCGCAGAGATCGTGAGCCTGACCGAGACCACCACCA 2081

Db   519 GGTCCAGAAGGAGCAGCCCGGCTTCTGGCACCAGATCCTCCCGAGATCCGGCGGGTGTCAG 578

Qy   2082 CCAGAAGACCGAGCTGCGCTGCAGCCATCCAGCTGGCCCTGCAGGACCAGGCCGCAGGTGAA 2141
           |||||||
Db   579 CGAGGAGACGACGACTGCCGTNATGAGCGTCGTGTACCAGCTGCACCGCCGCACGGCGAAGGCTGCT 638

Qy   2142 CATCGTGACCGACAGCCAGTACGCCCTGGGCCATCATCCAGGCCATCACCAGGACCAGGCGA 2201
           ||| || ||| |||| || ||
Db   639 GTTCCCGGCCGTCAACGTCAACGAC---TCTNTCACCAAGAGCAAGTTCGACAACATCTA 695

Qy   2202 GAGCGAGCTGGTGAACCAGATCATCGAGCAGTGCTGATCAAGAGAGAAGGTGTACCTGAG 2261
               ||| ||| ||||||| || ||| ||| | |
Db   696 CGGCTGCCGCCCACTCGCTCATCATCGACCAGTGGCGGCACCAAGCGGCGGCACCGACGTGATGCTCGGCGG 755

Qy   2262 CTGGGTGCCGCCCACCAAGGGCATCGGCGGCACCAGCATCGGCGCAACGTGGGCCAACAAGAGCTCGTGAGCAA 2321
           ||||||| ||| ||| | | |||
Db   756 CAAGGTCGCGCTGCTCGCCCGGCGTCGCTGCGCGGCCACGTGGGCCAAGGGCTGCGCCGGCATCGTGATCTACCAGTA 815

Qy   2322 GGGGATCGCCAAGGTGCGTGTTCCTGGACGGCCATCATGATGGCGGCGGCATCCTGATCCTGATCCTACCAGTA 2381
                                ||||| ||
Db   816 CGGCCAGGGCTGCGCCGGCGTCATCGTACGGCGAGATCGAGATCCAGCCCACCATCTGCGCGCTGCAGGCGGGTC 875

Qy   2382 CATGGACGACCTGTACGTGGGCAGCGACGGCGGCCCTAGGATCGA 2423
           |||||| ||| || | || ||
Db   876 GATGGCCCGCTTCGAGGTCAACACGCTCGAGGCGGGGGCTCGA 917

RESULT 3
CK159167/c
LOCUS   CK159167   869 bp   mRNA   linear   EST 05-DEC-2003
DEFINITION FGAS040564 Triticum aestivum FGAS: TALt5 Triticum aestivum cDNA,
            mRNA sequence.
ACCESSION   CK159167
VERSION   CK159167.1 GI:38985053
KEYWORDS   EST.
SOURCE   Triticum aestivum (bread wheat)
  ORGANISM Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 869)
  AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
            Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
            Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
            Penniket,C., Roach,J.L. and Sarhan,F.
  TITLE   Functional Genomics of Abiotic Stress In Wheat and Canola Crops
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Wm L Crosby
            Bioinformatics
            University of Saskatchewan, Department of Computer Science
            1C101 Engineering Building, 57 Campus Drive, Saskatoon,
            Saskatchewan, S7N 5A9, Canada
            Tel: 306 966 1769
            Fax: 306 966 2033
            Email: fgas_ests@cs.usask.ca
            This sequence is the direct result of the Base calling software
            Phred (default parameters). It is the raw base calls. To aid in the
            identification of the high quality insert the software Lucy identified
            (default parameters) has been run on this sequence. Lucy identified
            the region [128,636].
            Plate: TaLt537 row: N column: 23.
FEATURES   Location/Qualifiers
     source   1..869
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Wheat line PI 178383"
                     /db_xref="taxon:4565"
                     /lab_host="DH5 alpha"
                     /clone_lib="Triticum aestivum FGAS: TaLt5"
                     /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
                     subtractive hybridization) cDNA library from genotype
                     PI178383 cold hardened at 2 C for 21 days and 49 days

```
                    (equal amount of cDNA pooled together before subtraction,
                    tester) and subtracted against genotype Norstar cold
                    hardened at 2 C for 1 day (24 H)(driver) Modified Smart
                    cDNA (Clontech)priming and non-directional cloning"
ORIGIN

Query Match       3.3%;   Score 80.6;   DB 14;  Length 869;
Best Local Similarity  44.9%;  Pred. No. 2.7;  Indels  0;  Gaps  0;
Matches 305;  Conservative  0;  Mismatches  374;

Qy    580  AAGAGGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGACCATCATCGGCCGCAAC  639
Db    827  ATGGCGCGCGGCCCGACAACCAGGATCACCAACCAGGACGAGAACAACCACCACCACCAC  768

Qy    640  ATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCC  699
Db    767  CACAACGACGACAACAGCGGACACAACGACCACCAACAACAACAACAACAACAACAAC  708

Qy    700  GTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGCCCTACCACGAGGAG  759
Db    707  ACCAACAACAACAACAACAACAACAACGACAACCACCACCACCAACAACCACAACAAC  648

Qy    760  AAGATCAAGGGCCTGACCGCCATCTGCGAGGAGATGAGGAGGAGCAAGATCACCAAG  819
Db    647  AACAACAACGACAACAACAACAACAACAACAACGAGGAGGACCAGCACC  588

Qy    820  ATCGGCCCCGAGAGACCCCTACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACC  879
Db    587  AACAACAACAACAACAACAACGACAACAACAACAACAACAACAAC  528

Db    880  AAGTGCGCAAGCTGGCTTCCGCGGACTTCCGCGAGCTGAACAAGCGCACCAGACTTCTGGGAG  939
Db    527  AACAACAACAACAACAACAACAACAACAACAACAACAACAAC  468

Qy    940  GTGCAGCTGGGCATCCCCACCCCGCGGCCTGAAGAAGAGAGCGTGACCGTGCTG  999
Db    467  AACAACAACAACAACGACGACGACGACAACGACGACGACGACGAC  408

Qy   1000  GACGTGGGCGACGCCTACTTCAGCGTGCCCCTGACGAGGACTTCCGCAAGTACCACCGGC  1059
Db    407  GACAACAACGACAACAACGACAACAACGACAACAACAACAACAACAAC  348

Qy   1060  TTCACCATCCCCAGCATCAACAACGAGACCCCGGCATCCGCTACCAGTACACGTGCTG  1119
Db    347  GACCACAACGACAACGACAACGACAACAACAACAACAACAACCACAAC  288

Qy   1120  CCCCAGGGCTGGAAGGGCAGCCCCAGCATTCCAGAGCAGCATGACCAAGATCCTGGAG  1179
Db    287  AACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC  228

Qy   1180  CCCTTCCGCGGCCGCGAACCCCGAGATCGTGATCGTACCAGGCCCTGTACGTGGGCAGC  1239
Db    227  AACAACAACAACAACGACGACGACGTCAACGACGTGCCCAGTCAAGAACAACAACGACAAC  168

Qy   1240  GACCTGGAGATCGGCCAGC  1258
Db    167  AACGGCCAACAACAACAAC  149
```

RESULT 4
BM320864
LOCUS   BM320864   1132 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION   rockefeller.0.46 Mastigamoeba balamuthi lambda ZAP II Library
              Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
              sequence.
ACCESSION   BM320864
VERSION   BM320864.1   GI:18055270
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM   Mastigamoeba balamuthi
              Eukaryota; Pelobiontida, Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 1132)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,

```
              Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
              Philippe,H.
TITLE         The analysis of 100 genes supports the grouping of three highly
              divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL       Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE       21819461
PUBMED        11830664
COMMENT       Contact: Muller Miklos
              Laboratory of Biochemical Parasitology
              The Rockefeller University
              1230 York Avenue, New York, NY 10021, USA
              Email: mmuller@rockvax.rockefeller.edu
              Insert Length: 1132  Std Error: 0.00
              POLYA=No.
FEATURES             Location/Qualifiers
     source          1..1132
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match       3.3%;   Score 80.2;   DB 12;  Length 1132;
Best Local Similarity  45.3%;  Pred. No. 3.1;  Indels  9;  Gaps  2;
Matches 375;  Conservative  0;  Mismatches  444;

Qy    139  CGGCGCCCCCGCAAGAGAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGAC  198
Db     79  CGGCCGGAGGGCGAAGACGGACGGCGGCCCANCTGGTGATCGGCTTCCCCAGGGCCAAGGCC  138

Qy    199  TGCACCGAGCGCGCCAGGCCAACTTCTTCCGGAGGACCTGGACTTCCCAGGACAGGGCAAG----GTCCGGTTCACCACCAACAGCCGCGAGGGGGACATCGTC  192
Db    139  TACAACAGCCCCAAGTACCGCTTCGTC------GTCCGGTTCACCACCAACAGCGACATCGTC

Qy    259  CGCGAGTTCCCCAGCACGAGCAGAACCGCGCCAACAGCCCGCGGCGAGCTGCAGGTG  318
Db    193  TGCCAGATCGCTTACGCCAAGATCGACGACGACGCCGACCATCCTCGCCGCGCCGCCACTCGCAC  252

Qy    319  CGCGGGCGACAACCCCCGCAGGGAGGCGCCGGCGCCGCTGACCAACTACGCGCCTACGCGGACT  378
Db    253  GAGTCACCGCCTTCGGGGTCAGCTGCCCTCGTGGAGCATCCTCAAGGTGGGCGGCCAGATCAAGGAG  312

Qy    379  CAGATCACCCTGTGGCAGCGACCCCCTCGTGCTGAAGAAGGTCAACCTGACTTCCAAGTCAAGAGGT  438
Db    313  GGCCTGCTGCTGCGGCCCGCCGTGTGCTGAAGAAGCTCAACCTCGACTTCAAGTACGAGGGT  372

Qy    439  GCCCTGCTGACACCGGGCGCCGACGAACAGAATGAGCGTCGAGGATGAGCGACGGGGCCCGCGT  498
Db    373  GTCAAGAAGGTCAACGGCTCAACGGCCGAGGACTACAACGTCGGCGGCTTCATCAAGGTGGCCCAGTACGACCAG  432

Qy    499  TGGAAGCCCAAGATGATCGGCGGCATCGGCGTTCGGCCTTCGGACCTCCGACACCTGGACCGTCGACCGTGTTCGCC  558
Db    433  TTCAAGGGCCTGCTTGACGTCGGCCTGCTCGTGCTCCACCTCACCACTGCCCC---CACAGCGAGACCCGGTTC  492

Qy    559  ATCCTGATCGAGATCTCGCGGAAGAGCCAGCTGCTGCTGCCCACCCCC  618
Db    550  GTCGGCTTCAACGGCGACAAGAGGAGCTCAACGGCGGGTTCCCGCAAGTACATCTTC  609

Qy    619  GTGAACATCATCGGCGCCGCAACATGCTGACCCAGCTGGCTGCACCCTGAACTTCCCCATC  678
Db    679  AGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGGCATGCATGGACGAGCAGGTGAAG  738

Qy    739  CAGTGGCCCCTGACCGCGGAGAAGATCAAGGGCCCATCTGCGAGGAGATGGAG  798
Db    670  CGGCGGCCACGTCGCCGCTACGCCAAGGAGGGTGTCACCGCCATCTGCTCGAGAAGATCTAC  729

Qy    799  AAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACAACACCCCCGTGTTCGCC  858
```

RESULT 5

LOCUS   BM320900   1165 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION   rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
             Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
             sequence.
ACCESSION   BM320900
VERSION   BM320900.1 GI:18055306
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM   Mastigamoeba balamuthi
             Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 1165)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
             Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
             Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
             divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
   MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
             Laboratory of Biochemical Parasitology
             The Rockefeller University
             1230 York Avenue, New York, NY 10021, USA
             Email: mmuller@rockvax.rockefeller.edu
             Insert Length: 1165   Std Error: 0.00
             POLYA=No.
FEATURES   Location/Qualifiers
     source   1..1165
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.28%;   Score 79.6; DB 12; Length 1165;
Best Local Similarity 45.1%;   Pred. No. 3.7;
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;

Qy   90 CATCAAGTGCTTCAACTGCGCGCAAGGAGGGCCACATCGCCCGCAACTGCGCGCCCCCCG 149
              ||||||||| ||||||||| || |||| ||| || |||||| ||| ||||||
Db   46 CGTCAAGAACAAGGCGTACTTCAAGCGCTTCAAGCGCTTCCGTCGCTCCGTCGCCGAGGG 105

Qy   150 CAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGAGATCAAGGTGCG 209
              |||| || | |||| ||||||||| ||| |||||||||| || ||
Db   106 CAAGACGCGACTACCCGCGCGCCGCCCANCTGGTCATCCAGGACAACAAGCAACAGCCCC 165

Qy   210 CCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCACCCCAGGGCAAGGCCCGGCAGTTCC 269
              | || ||| | | | | | || ||| || |
Db   166 CAAGTACCGCTTCGTC------GTCCGCTTCACCAACGCGGGACATCGTCTGCCAGATCGC 219

Qy   270 CAGCGAGCAGAGAACCGCGCAGGACGCGGCGACCACCATCCTCGCCGCGGCAGGGCGACAA 329
              || || | | || |||||| || |||| | | || | | || ||| |
Db   220 CTACGCCAAGATCGACGCGGCCGGCGACGCCATCCTCGCCTACTCTGCAGAGCTCACCCG 279

Qy   330 CTTCGGCGTCAAGTGCTTCGGCGTCCGGCTGACCAACTACGCGCCGCCTACGCGACTGGC 389
              ||||| |||| | | || | | || || | ||||| |||||| ||
Db   280 CCCCCGCAGCGAGCGGCCGCGGGCGCCGAGCGGCGCCACGGGGCGCCCGGGCGCTGCTGCT 339

RESULT 6

LOCUS   BM321430   867 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION   rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library
             Mastigamoeba balamuthi cDNA similar to ribosomal protein S4, mRNA
             sequence.
ACCESSION   BM321430
VERSION   BM321430.1 GI:18055836
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM   Mastigamoeba balamuthi
             Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 867)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
             Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
             Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
             divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
   MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
             Laboratory of Biochemical Parasitology
             The Rockefeller University
             1230 York Avenue, New York, NY 10021, USA
             Email: mmuller@rockvax.rockefeller.edu
             Insert Length: 867   Std Error: 0.00

SOURCE   Mastigamoeba balamuthi
ORGANISM   Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 1550)
AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
TITLE   The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE   21819461
PUBMED   11830664
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, New York, 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 1550 Std Error: 0.00
            POLYA=No.
FEATURES   Location/Qualifiers
     source   1..1550
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

Query Match   3.1%; Score 76.6; DB 12; Length 1550;
Best Local Similarity 45.6%; Pred. No. 9;
Matches 308; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

Qy 1749 ...
Db   5 ...

[sequence alignment data]

RESULT 7
BM321022
LOCUS   BM321022   1550 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller_0.1192 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION   BM321022
VERSION   BM321022.1 GI:18055428
KEYWORDS   EST.

```
Qy   2349  CGGCATCGATGGGCGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGG  2408
                  ||||||   |||||||||  ||| ||||||||||  |||  ||||||||||||||
Db    602  GGAGATCGCACCCCATCTGCGCGGTCGAGGCGTCGATGGCGGCCTTCGAGGTCAACACGCT   661

Qy   2409  CGGCCCTAGGATCGA  2423
                  ||  ||||||||||
Db    662  CGAGGCGGGGCTCGA   676
```

```
RESULT 8
BI724851
LOCUS          BI724851                     545 bp    mRNA    linear    EST 19-SEP-2001
DEFINITION     1031075E05.yl C. reinhardtii CC-1690, Stress II (normalized),
               Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      BI724851
VERSION        BI724851.1    GI:15700546
KEYWORDS       EST.
SOURCE         Chlamydomonas reinhardtii
  ORGANISM     Chlamydomonas reinhardtii
               Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
               Chlamydomonadaceae; Chlamydomonas.
REFERENCE      1  (bases 1 to 545)
  AUTHORS      Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
               Lefebvre,P., McDermott,J.P., Shrager,J., Siflow,C. and Stern,D.
  TITLE        Analyses of the Chlamydomonas reinhardtii Genome: A Model,
               Unicellular System for Analyzing Gene Function and Regulation in
               Vascular Plants. Project: 1031
  JOURNAL      Unpublished (2001)
COMMENT        Contact: Charles Hauser
               DCMB Box 91000
               Duke University
               Durham, NC 27708-1000
               Tel: 919 613 8159
               Fax: 919 613 8177
               Email: chauser@duke.edu.
FEATURES             Location/Qualifiers
     source          1..545
                     /organism="Chlamydomonas reinhardtii"
                     /mol_type="mRNA"
                     /strain="CC-1690 wild type mt+ 21gr"
                     /db_xref="taxon:3055"
                     /clone_lib="C. reinhardtii CC-1690, Stress II
                     (normalized), Lambda Zap II"
                     /note=Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
                     XhoI. Stress condition II library, constructed by John
                     Davies and Jeffrey McDermott, combines cDNAs from CC-1690
                     cells grown to mid-log phase in TAP (NH4+ - containing)
                     and shifted to TAP - NO3- (24hrs); H2 production
                     conditions (0, 12hr, 24hr) see Melis et al.,(2000) Plant
                     Phys. 122; 127-135; TAP+ H2O2 (1, 12, 24 hr); TAP +
                     sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
                     PolyA mRNA was purified from each sample, pooled and cDNA
                     synthesized. The cDNA was directionally cloned into lambda
                     Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
                     sites. pBluescript II SK- plasmids were excised from the
                     lambda ZAP clones by superinfection with ExAssist
                     (Stratagene) phage. The library was normalized using
                     method 4 described in Bonaldo et al., (1996) Genome
                     Research 6: 791-806."
ORIGIN
Query Match        3.1%;     Score 75;   DB 12;   Length 545;
Best Local Similarity   46.7%;   Pred. No. 11;
Matches   237;    Conservative    0;  Mismatches  270;  Indels  0;  Gaps  0;
```

```
Qy   1184  TCCGCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCAGCGACC  1243
                  |||||||   |||   ||||| |||||    |||||||       ||||||||||||
Db     36  TGCGGGCGCAGCAGCTGTGTGTTTGCCATCCTGGAGGAGAGCGGCCCCAGGGCAACCCCT    95

Qy   1244  TGGAGATCGGCGACACCGGCGCCAAGATCGAGGAGAGTCGCGCAAGCACCTGCTCGCGTGGG  1303
                  |||  | ||   | |||||    ||  |||||||||    ||||||||| |||||||
Db     96  ACATGCTGTACAAGGACCACTGCAACCGAGATCCAACCAGCAGAACCTGGCACCATCA   155
```

```
Qy   1304  GCTTCACCACCCCGACACAGAAGCACCAGAAGAGGAGCCCCCTTCTGCCATCGAGCTGC  1363
                  |||   |||||   |||||| ||||||||   ||||||||||||||||||||||||
Db     156  AGTGCTCCAACCTGTGCACGGAGATTATCGAGTACACCAGCCCCGATGAGATCGCCGTGT   215

Qy   1364  ACCCCGACAGTGGACCATGCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGA  1423
                  |||||   | |||||| ||||||||||||  |||||  ||||||||||||||||
Db     216  GCAACCTGGCCTCCATCGCCGCTGCCCGGCTTCGTGCGGGAGGGCCAGCCCACCGGCT   275

Qy   1424  ACGACATCCAGAAGCTGGTGCGCAAGCTGAACTGGGCACGCGGCCCAGATCTACCCCGGCATCA  1483
                  ||    |||||||  || |||||||||||  ||||||||||||||||   |||||||||
Db     276  GGGAGAGCAAGAAGCTGGTGGCTCGTGCGGCGCCAAGGCCCAACCGCGTTCTTCGACTTCGAGA   335

Qy   1484  AGGTGCGCCAGTGTGCAAGCTGCTGCCGCGGCCCAAGGCCCTGACGACATCGTGCCCC  1543
                  ||||||    |||  |||||||||||||||   |||||||||||||||| ||||||||
Db     336  AGCTGGCTGGATGACCAAGATGGTGACCCGCCAACCTGACCAACAAGATCATTGACGTCAACT   395

Qy   1544  TGACCGAGGAGGCCGAGCTGGAGCTGGCCGGCAAGCTGGCCGAGACCGAGAACCGCGAGATCGTGGCGGAGCCCGTGC  1603
                  |||||||||||  ||||||| |||   ||||||||||||||||||||  ||||||||||| |||||||||||||||
Db     396  ACTACCCGCACCCCAAGGCGCAGCGCTCCAACATGCGCCCACCCCCATCCGGCCTGGGCG   455

Qy   1604  ACGGGCGTGTACTAGGACCCCAGCAAGGACCTGGTGCGGGAGATCCAGAAGCAGGGCCACG  1663
                  |||||||||||||     | |||||||| |||||||  |||||||||||||||||||||
Db     456  TGCAGGGCCTGGCGGACACCTTCATCCTGTGGGCATGGCCTTCGACTCGCCCGAGGCCG   515

Qy   1664  ACCAGTGGACCTACCAGATCTACCAGG  1690
                  |||||||| |||||||||||||||||
Db     516  CCGAGCTGAACCGCCCAGATCTTCGAGG   542
```

```
RESULT 9
BM321023
LOCUS          BM321023                     862 bp    mRNA    linear    EST 03-JAN-2002
DEFINITION     rockefeller.0.594 Mastigamoeba balamuthi lambda ZAP II Library
               Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
               3.3.1.1), mRNA sequence.
ACCESSION      BM321023
VERSION        BM321023.1    GI:18055429
KEYWORDS       EST.
SOURCE         Mastigamoeba balamuthi
  ORGANISM     Mastigamoeba balamuthi
               Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE      1  (bases 1 to 862)
  AUTHORS      Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
               Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
               Philippe,H.
  TITLE        The analysis of 100 genes supports the grouping of three highly
               divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  MEDLINE      21819461
  PUBMED       11830664
COMMENT        Contact: Muller Miklos
               Laboratory of Biochemical Parasitology
               The Rockefeller University
               1230 York Avenue, New York, NY 10021, USA
               Email: mmuller@rockvax.rockefeller.edu
               Insert Length: 862  Std Error: 0.00
               POLYA=No.
FEATURES             Location/Qualifiers
     source          1..862
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN
Query Match        3.1%;     Score 75;   DB 12;   Length 862;
Best Local Similarity   46.2%;   Pred. No. 12;
Matches   283;    Conservative    0;  Mismatches  327;  Indels  3;  Gaps  1;
```

```
Qy   1722  CGCCAAGATGCGCACCCGCCCCACCACCAACGACGTGAAGCACGTGACCGGCCGTGCAGAA  1781
                  ||||   |||    ||||||||    |||||||||||||||||||||   |||
```

/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_7463"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN

Query Match　　　　3.0%;　Score 74.4;　DB 28;　Length 914;
Best Local Similarity　47.0%;　Pred. No. 15;
Matches　265;　Conservative　0;　Mismatches　296;　Indels　3;　Gaps　1;

```
Qy    652  CTGGGCTGCACCTGAACTTCCCATCAGCCCATCCAGACCGTGCCGTGAAGCTGAAG  711
                ||||  ||||||||||  |||||  |||||   |||||| |||||||
Db    112  CTGCGCAACCAGCAGGACGGCCTGATCTTCAGCCAGGCGGCGCCCTCGAGGTGCAGGCC  171

Qy    712  CCCGGCATGGACGGCGCCCCAAGGTGAAGCAGTGGCCCCTGACCGGAGGAGAGATCAAGGCC  771
                |||||  |||||||||  ||||||  |||||||| |||||||||||
Db    172  GGCAGCCTGGACAACCGCCAGGACATCCTCCCGAGATCCCGGTATC  231

Qy    772  CTGACCGGCCATCTCGGAGGAGGATCGGAGGAGGGCAAGATCACCAAGATCGGCCCGAG  831
                ||||||||  |||||  |||||||||  ||||||||| |||||||||
Db    232  GGTGGCGCTGGACAACCAGGGGCCGGCCCTGGACAGCGGGCCGGTAACCTAGACCTG  291

Qy    832  AACCCTACAAACACCCCCGTGTTCGCCATCAAGAAGAAGAGCAGCACAAGTGGCGCAAG  891
                ||||||| |||  |||||  |||||  ||||||||| |||||||
Db    292  CAGAGCGGCAGCCTGACAACGGCGCCCCTGCTCAACAGCGCCAAGGGTTGGCTG  351

Qy    892  CTGGTGGACTTCCGGCAGCTGACAAGCGCACCCGAGACTTCTGGGAGGTGCAGCTGCGGC  951
                ||||||| |||||  |||||||||  ||||||  ||||| ||||||||
Db    352  AAGCTGGTCACCGGCGTGTTGACAACAGCGCCGCGCAACCAGCGCAGTCGCTGGAA  411

Qy    952  ATCCCCACCCCCGGGCCTGGAGGAAGAAGAGCCGTGACCGTGCCTGACGTGGGGCGAC  1011
                ||||||  |||||||||  |||||||  ||||||||||  |||
Db    412  ATCCGGCGCCGGCAAGGCGTGGGCGCAACCAGCAGGGC---CATGTCTCCGCGCGGCGGC  468

Qy    1012 GCCTACTTCAGGGTGCCCCTGGACGGAGACTTCGCGCAAGTACACCAGGGTGGGGCCTCC  1071
                |||||||  ||||||  |||||||||  ||||||||||  ||||||
Db    469  GACAACCGCATCGTCACCGTCGACTTCGACAACCAGGGTACCAACGTGCTGCCCGGCGGC  528

Qy    1072 AGCATCAACAACGAGACCCCCGGCATCCTACCAGTACAACGTGCCCGGCGGGCTGG  1131
                |||||||  ||||||  ||||  ||||||||  ||||||
Db    529  CTGCTCAGCCTGCAGCCGGCGGCCCCAGGGCGCGGCGGCCAGGGGCGC  588

Qy    1132 AAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGCCCTTCCGCGC  1191
                 ||||||  |||||||||||| ||||||  ||||||||||||  |||
Db    589  AAGTTCGGGCCCACCATCGAACTTCAGCCTCGGCCGGCGCTGGCCAACCGCTTCGC  648

Qy    1192 CGCAACCCGAGATCGTGATCTAC  1215
                |||||||||||||||||
Db    649  AGTTGGAAAGCGAGCGAGCTGC  672
```

RESULT 11
AY401196
LOCUS　　　　AY401196　　　　　　2433 bp　　DNA　　　linear　　GSS 12-DEC-2003
DEFINITION　Homo sapiens HCM0810 gene, VIRTUAL TRANSCRIPT, partial sequence,
　　　　　　　genomic survey sequence.
ACCESSION　AY401196
VERSION　　AY401196.1　GI:39757185
KEYWORDS　GSS.
SOURCE　　Homo sapiens (human)
　ORGANISM　Homo sapiens
　　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　　Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE　1　(bases 1 to 2433)
　AUTHORS　Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
　　　　　Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
　　　　　Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
　　　　　Adams,M.D. and Cargill,M.
　TITLE　Inferring nonneutral evolution from human-chimp-mouse orthologous
　　　　gene trios
　JOURNAL　Science 302 (5652), 1960-1963 (2003)

```
Db    220  CGCCAAGATCCGCTGGTGTCGTCGTGCAACATCTTCTGCGACGCGGGCGGCCCGGC  279

Qy    1782 GATCGCCATGAGCAGGACGATCGTGATCTGGGGCAAGAGTCCGGCTGCCATCCA  1841
                |||||||  ||||||  |||||  ||||||||
Db    280  CATCGCGCGCAGGCCGGAGTCTCGGTCTTCGCCTGGAAGGGCGGAGACCTTCCAGGAGTACTG  339

Qy    1842 GAAGGAGACCTGGGAGACCTGGTGACCACTGGCAGCCGACTACTGGGCAGCCCACCTCGATCCCCGAGTG  1901
                |||||  |||  ||||  ||||  |||||||
Db    340  GGAGTGCACCTNGAAGGCCCTGTGCTTGCCTCGGCCCTTACCAGGGACCTCAGANCATCGTCGA  399

Qy    1902 GGAGTTCGTGAACACCCCCCCGGTGAAGCTGGTGTACCAGCTGGAGAAGGAGCCCAT  1961
                ||||||  |||||||  |||  |||||||||  |||||
Db    400  CGACGGCGGTGACGCGACTCTAATGATCCACAAGGGGTTCGCGGCCAACCGAGACCAAGATCGGCAA  459

Qy    1962 CATCGGCGGCCGAGACCTTCTACGTGGACGGCGCCGCCCGCGAGACCAAGATCGGCAA  2021
                ||||||||  ||||||||  ||||||||||  ||||
Db    460  GCTGCTGGAGGACGACGAGGGGCCTCGAGGAGTGGCCGCTGCTCAACAACGTGCTCAAGCA  519

Qy    2022 GGCCGGCTACGTGACCGACGGCCGGCGGCCGCAGAAGATCGTGACCTGACCGGACCACCA  2081
                |||||  ||||||||||  |||||||||  |||||
Db    520  GGTCCAGAAGGAGCAGCTGCAGGCCATCCAGTTCTGCCACAAGAGGCCTCCCCGAGATCCCGGTGTCAG  579

Qy    2082 CCAGAAGACCGAGCTGCAGGCTGCCATCCAGCTGGCCTGCCGGGACCAGCGGCAGCGAGGTGAA  2141
                ||||||||||  |||||||  |||||||||||  |||||
Db    580  CGAGGAGACCGACGCGACTGGCGTGATGAGGCTGTACCAGCTGACCTGCCGCGCAGCGGCAAGCTGCT  639

Qy    2142 CATCGTGACCGACAGCCAGTACGCCCTGGGCATC---ATCCAGGCCCGACAAGAG  2198
                |||||||||||||||  |||||  |||||  |||||
Db    640  GTTCCCCGGCCGTCAACGTCGTCACCAGCGACTCTGTCACCAAGAGCAAGTTTGACACATCTACGG  699

Qy    2199 CGAGAGCGAGTGGTGTAGAACCAGATCATCGAGCAGATCATCAAGAAGGAGAAGGTGTACCT  2258
                ||||||||||  |||||||  ||||||  ||||||||||
Db    700  CTGCGCCACTGCTCGCCATCGAGATCGGCCGGCATCAAGCGCGGCGCCGACGTGATGCTCGGCGCGCAA  759

Qy    2259 GAGCTGGTGCCCGCCCACAAGGGCATCGGCGGCGCAACGACGCAGATCGACAAGCTGGTGAG  2318
                ||||||  |||||||||  |||||||||  ||||||
Db    760  GGTCGCGTCGTCGTCGGACGTGGGCAAGGGCGCTCACGCGGCGGCGCGGGCGCGG  819

Qy    2319 CAAGGGCATCCGC  2331
                ||||||||||||
Db    820  CCAGGGCTGCCCGC  832
```

RESULT 10
BZ568300
LOCUS　　　　BZ568300　　　　914 bp　　DNA　　　linear　　GSS 17-DEC-2002
DEFINITION　pacs2-164_7463.y2 pacs2-164_7463, genomic survey sequence.
ACCESSION　BZ568300
VERSION　　BZ568300.1　GI:27201058
KEYWORDS　GSS.
SOURCE　　Pseudomonas aeruginosa
　ORGANISM　Pseudomonas aeruginosa
　　　　　　Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
　　　　　　Pseudomonadaceae; Pseudomonas.
REFERENCE　1　(bases 1 to 914)
　AUTHORS　Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
　　　　　Burns,J.L., Kaul,R. and Olsen,M.V.
　TITLE　Whole-Genome-sequence variation among multiple isolates of
　　　　Pseudomonas aeruginosa library
　JOURNAL　J. Bacteriol. (2002) In press
COMMENT　Contact: Chris K. Raymond
　　　　　Genome Center
　　　　　University of Washington
　　　　　Box 352145, Seattle, WA 98105-2145, USA
　　　　　Tel: 2062216954
　　　　　Fax: 2066857244
　　　　　Email: craymond@u.washington.edu
　　　　　Class: shotgun.
FEATURES　　　　　Location/Qualifiers
　source　　　　1..914
　　　　　　　　/organism="Pseudomonas aeruginosa"
　　　　　　　　/mol_type="genomic DNA"

```

```
  PUBMED   14671302
REFERENCE   2 (bases 1 to 2433)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
     source   1..2433
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>2433
                     /locus_tag="HCM0810"
ORIGIN

  Query Match   3.0%; Score 73.8; DB 29; Length 2433;
  Best Local Similarity   43.0%; Pred. No. 21;
  Matches 485; Conservative   0; Mismatches 632; Indels   12; Gaps   2;

Qy   463 GACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGC 522
              |||| | | | || | | | | | || || |||| || || | |
Db   1240 GACAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTC 1299

Qy   523 ATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAG 582
              || | | |||| | || |||| | | | | ||| || | |
Db   1300 ATGCTCAGTGGCATCCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCCTCAAGCTG 1359

Qy   583 AAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGAACATCATCGGCCGCAACATG 642
               || ||| | |||| ||| |||| | | | ||| |||| ||| |
Db   1360 GAGCTGATCCCCGACGTGACCATCCCGCCCAGCATTGCCCAGCTCACGGGCCTCAAGGAG 1419

Qy   643 CTG-------ACCCAGCTGGGCTGCACCCTGAACTTCCCCATCAG-----CCCCATCGAG 690
              ||| ||| | | | | |||| ||| | | || ||||
Db   1420 CTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCCGCGCTGGCCTTCCTGCGCGAG 1479

Qy   691 ACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGGTGAAGCAGTGGCCCCTG 750
              | | ||| | | || | || ||| | || | || |||| |
Db   1480 AACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTAT 1539

Qy   751 ACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGGAGGGCAAG 810
              | | | ||| | ||| ||| | | || | || | | || | | |
Db   1540 AGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGC 1599

Qy   811 ATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGTTCGCCATCAAGAAGAAG 870
                || | |||| | | ||| || | | | ||||
Db   1600 TACATCGTCATCGACGGGCTGCGGGAGCTCAAACGCCTCAAGGTGCTGCGGCTCAAGAGC 1659

Qy   871 GACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGAC 930
               || | |||| || || |||| | || || || | | | |
Db   1660 AACCTAAGCAAGCTGCCACAGGTGGTCACAGATGTGGGCGTGCACCTGCAGAAGCTGTCC 1719

Qy   931 TTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTG 990
               || | | | || | || | | |||| ||||||| | || | ||
Db   1720 ATCAACAATGAGGGCACCAAGCTCATCGTCCTCAACAGCCTCAAGAAGATGGCGAACCTG 1779

Qy   991 ACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTGCCCCTGGACGAGGACTTCCGCAAG 1050
              || | ||||| || | | || | || || |||| ||
Db   1780 ACTGAGCTGGAGCTGATCCGCTGTGACCTGGAGCGCATCCCCCACTCCATCTTCAGCCTC 1839

Qy   1051 TACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTAC 1110
               ||| | || | || ||||| | | |||| | |
Db   1840 CACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATC 1899

Qy   1111 AACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAG 1170
              | | | | || || | | | | || | | | | | | | ||| || |
Db   1900 AGCTTCCAGCACCTGCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTAC 1959

Qy   1171 ATCCTGGAGCCCTTCCGCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTAC 1230
              |||| | || || || || | | | |||||| | || | |
Db   1960 ATCCCCATCCAGATCGGCAACCTCACCAACCTGGAGCGCCTCTACCTGAACCGCAACAAG 2019

Qy   1231 GTGGGCAGCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCAC 1290
               | | | || | || || |||| | | ||| | |
Db   2020 ATCGAGAAGATCCCCACCCAGCTCTTCTACTGCCGCAAGCTGCGCTACCTGGACCTCAGC 2079

Qy   1291 CTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTG 1350
              | | | |||| | | |||||| | || | || | | || |||
Db   2080 CACAACAACCTGACCTTCCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCTA 2139

Qy   1351 CCCATCGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAG 1410
               ||||| | ||| | || | | | | || | | | |||| |
Db   2140 GCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTG 2199

Qy   1411 AGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATC 1470
               | | || | | |||| | || | ||| || || |
Db   2200 CGGGCCCTGCACCTGGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGGTGGGCGAGCTG 2259

Qy   1471 TACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACC 1530
                | | | || | | || | |||| | |||| || | |
Db   2260 ACCAACCTGACGCAGATCGAGCTGCGGGGCAACCGGCTGGAGTGCCTGCCTGTGGAGCTG 2319

Qy   1531 GACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACC 1579
              | | || ||| | || || ||| | ||| ||| |||
Db   2320 GGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGAGGACC 2368

RESULT 12
BM321393
LOCUS   BM321393   853 bp   mRNA   linear   EST 03-JAN-2002
DEFINITION rockefeller.0.1222 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
            sequence.
ACCESSION   BM321393
VERSION   BM321393.1 GI:18055799
KEYWORDS   EST.
SOURCE   Mastigamoeba balamuthi
  ORGANISM Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 853)
  AUTHORS   Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
  TITLE   The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  MEDLINE   21819461
   PUBMED   11830664
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 853   Std Error: 0.00
            POLYA=No.
FEATURES   Location/Qualifiers
     source   1..853
                     /organism="Mastigamoeba balamuthi"
                     /mol_type="mRNA"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"
ORIGIN

  Query Match   3.0%; Score 73.2; DB 12; Length 853;
  Best Local Similarity   46.9%; Pred. No. 20;
  Matches 299; Conservative   0; Mismatches 329; Indels   9; Gaps   2;

Qy   388 CTGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTG 447
              ||| | || ||| || |||| | | | | || | ||| | | | |
Db   125 CTGGCCCGCCGTGTGCTGAAGAAGCTCAACCTCGACTCCAAGTACGAGGGTGTCAAGAAG 184
```

                                           chromosome)"
                                           /db_xref="taxon:7165"
                                           /clone="196004969611O"
                                           /dev_stage="Adult"
                                           /lab_host="DH10b"
                                           /clone_lib="A.Gam.ad.cDNA.blood1"
                                           /note="Vector: pSport; Site 1: SalI; Site 2: NotI; Whole
                                           adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
                                           hours after human blood feeding. cDNA inserts >500 bp
                                           cloned directionally into pSport 1. Not 1 site is 3'.
                                           Clones available through the Malaria Research and
                                           Reference Reagent Resource Center (www.malaria.mr4.org)"

ORIGIN

Query Match            3.0%;    Score 72.8;   DB 12;   Length 566;
Best Local Similarity 48.3%;    Pred. No. 21;
Matches 203;   Conservative  0;   Mismatches 217;   Indels  0;   Gaps  0;

Qy   1944 GCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCACGTGGACGGCGCCCAACCG 2003
               ||||| ||   |   ||||||   || || || | |||||| ||||||  ||
Db    482 GTTCGAGATCGACGCCAACGGCATCCTGCAGGTGTCGGCGCCGAGGACAAGGCAA  423

Qy   2004 CGAGACCAAGATCGGCAAGCCGGCTACGTGACCACCGGGCGCCGCAGAAGATCGTGAG 2063
               |||| ||| || || |||||||  || |  |||| ||| |||||  ||| ||
Db    422 CCGGAAAGATCGTCATCACCACCAGCACCAGAACCGCTGACGCCGACATCGAGCG  363

Qy   2064 CCTGACCGGACCACCACCAACAGAGAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGA 2123
                ||||| || ||  || ||| ||||  ||||| || ||||| ||  ||
Db    362 CATGATCAAGGATGCGGGAGCGGTTCGCCGACGACGACAAGAAGCTGAAGGAGCGCGGTGGA  303

Qy   2124 CAGCCGCAGGCAGCTGAACATCGTGACCGACGACGACCAGCAGCCGCCCTGGGCATCATCCAGGC 2183
               || |||||| |||||   |||| |||| || || ||||| ||||||  || ||
Db    302 GGCCCGCAACGAGCTCGAGTCGAGAGCTGACAGCGTACAGGCTGAAGAACCAGCTCAGCTCGAAGGA  243

Qy   2184 CCAGCCCGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGTCGATCAAGAA 2243
               ||||||| || ||| ||| ||| ||  ||| |||| || || ||| || ||  ||
Db    242 CAAGCTGGGCGGCGAGCGTGTCGGACGACGAAGATGGAGGAGCGATCGACGA  183

Qy   2244 GGAGAAGGTGTACTTGAGCTGGGTGCCCGCCCCACAAGGGCGCATCGGCATCGATGGCGG 2303
               ||| |||||| ||  ||||||| ||||| |||||| |||| || || ||  || ||
Db    182 GAAGATCAAGTGGCTGGACGAGAACCAGGACACCGAGGCGGAGTACAAGAAGCAGAA  123

Qy   2304 CGACAAGCTGGTGAGCAAGGGCATCCGCAAGGGTCGTGTTCCTGGACGGCATCGGATGGCGG 2363
               ||| |||||  |||| ||||| || || ||||| || ||| || ||||| || ||
Db    122 GAAGGAGCTCGGAGGACAGCATCCTGCAGCCCATCATTGCCAAGCTCGTTACGCGCAGTGGGCGG  63
```

RESULT 14
CB643171

```
LOCUS       CB643171                 788 bp    mRNA    linear   EST 08-APR-2003
DEFINITION  OSJNEb03L13.f OSJNEb Oryza sativa (japonica cultivar-group) cDNA
            clone OSJNEb03L13 5', mRNA sequence.
ACCESSION   CB643171
VERSION     CB643171.1  GI:29638162
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 788)
  AUTHORS   Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  TITLE     Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
```

```
Qy    448 GACACCGGCGCCGACGACACCGGTGCTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAGCCC 507
               ||||||| ||  |   || ||  ||  | ||||||||  ||  || |||| ||||||||
Db    185 GTCAACGGCGAGGACTACAACGTCGAGGAGCTCGAGGAGCTCGAGGAGCTCGAGGAGC 244

Qy    508 AAGATGATCGGCGGCATCGGCGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATC 567
               ||   |||| ||| ||||| || || || ||  |||| | || || ||  ||||
Db    245 CTGCTCGAGTCGGCCTGGTCCGACCTCGACCTCGAGGCCCGCGGTGTTCGCGCCCCTCAAG 304

Qy    568 GAGATCTGCGCAAGGAGAGGCCATCGGCACCGTGCTGATCGGCGCCCACCCCGTGAACATC 627
               | |  |||| ||| | ||| || ||| ||  | ||| || || |||| || || ||
Db    305 GGCAATGTCGACGGCGGCGGCGTCAACGTCCCC---CACAGCGAGACCCCGTCGTCGGCTTC 361

Qy    628 ATCGGCCGCAACATGCTGACCCGACCTGGGCTGCGCACCCTGAACTTCCCCATCAGCCCCATC 687
               || ||| || || ||||  || || || |||| |  ||| ||  ||  || ||||
Db    362 AACGGGCGACAAGAAGGAGCTCATCGCGGGTCCTCCGCAGTTACACTTCGGCCGGCCAC 421

Qy    688 GAGACCGTGCCCGTGAAGCTGAAGCTGGAGCCGGCGATGGCGGCCCCAAGTGAAGCAGTGGCCC 747
               || |||| |||  ||| || || || | || || |||||| ||| || |  || |
Db    422 GTCGCCGGTCACATGAGCTCCTCAAGAGCACCGCCGCCGCCGCCGTTCAAGCCGCCAGTTC 481

Qy    748 CTGACCGAGGAGAAGATCAAGGCCCTGACCGCCCATCTGCAGCCGCCATCTGCGAGGAGATGGGGC 807
               || || |||| ||| ||  ||| || || || | |||||| || || || ||||||||| ||
Db    482 TCGCGCTACGCCAAGGAGGGTGTCACCGGCGACATGCTGCAGAAGATCTACACCGAGGGCC 541

Qy    808 AAGATCACCAAGATCGGCGCCCGAGAACCCCTACAACACCCCGTGTTCGCCATCAAGAAG 867
               || || || |||||| || || ||  ||| |||| || |||  || ||| ||||
Db    542 CA------CAAGCAGATCCGCGCCGACCCGAACCTTGTCCCCAAGGCGCCTCGAAGCCC 59?

Qy    868 AAGGACAGCACCAAAGTGCGCGAAGTGCTGCAAGTTCCGCGAGCTGACTTCGACAGCCGCCACCCAG 927
               || ||  ||  || ||  || ||  || || ||   || ||| || || | ||||||| || ||||
Db    596 GAGGGCGCCAAGCCCAGACACTGGGGCAAGCAGCGCGGCAGCGTGACCTACCACGAGCGCCAAGAAC 655

Qy    928 GACTTCTGGGAGTGCAGCTGGCATCCCCACCCCCGCGGGCCTGAAGAAGAAGAGAGAGAGC 987
               ||| || || || |||| ||| || |||  || |||||||| ||| || || |||||| ||
Db    656 CGCGTGCGCGAGAAGAGGTCCGCCTGGCTTACCCGACGCCCCAAGAGCAGCCAGTAAATT 715

Qy    988 GTGACCGTGCTGGACGTGGGGCCGACCGCCTACTTCAGCG 1024
               ||||| ||  | || || |||  ||| || || || ||||
Db    716 CCCGGGTGCCCTGTACACNGCCATCTCCCCGCTGCGGCG 752
```

RESULT 13
BM587428/c

```
LOCUS       BM587428                 566 bp    mRNA    linear   EST 25-FEB-2002
DEFINITION  1700068732I202 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone
            196004969611O 5', mRNA sequence.
ACCESSION   BM587428
VERSION     BM587428.1  GI:18883289
KEYWORDS    EST.
SOURCE      Anopheles gambiae (African malaria mosquito)
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1  (bases 1 to 566)
  AUTHORS   Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L.,
            Charlab,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
  TITLE     Celera Anopheles gambiae EST project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Holt R.A.
            Celera Genomics
            45 W. Gude Dr., Rockville, MD 20850, USA
            Tel: 2404533151
            Fax: 2404534580
            Email: HoltRA@celera.com
            Plate: NU01004AAX row: G column: 08
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..566
                     /organism="Anopheles gambiae"
                     /mol_type="mRNA"
                     /strain="RSP-ST (Reduced susc. to Permethrin - std.
```

```
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: gta aaa cga cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 03  row: L  column: 13
Seq primer: gta aaa cga cga cgg cca gtg.
FEATURES
    source          Location/Qualifiers
                    1..788
                    /organism="Oryza sativa (japonica cultivar-group)"
                    /mol_type="mRNA"
                    /cultivar="Nipponbare"
                    /db_xref="taxon:39947"
                    /clone="OSJNEb03L13"
                    /tissue_type="Leaf"
                    /dev_stage="3 week"
                    /lab_host="DH10B"
                    /clone_lib="OSJNEb"
                    /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                    XhoI; 24 hrs after innoculation with Rice Blast (Che
                    8606l)"
ORIGIN

Query Match       3.0%;    Score 72.6;   DB 14;   Length 788;
Best Local Similarity   46.3%;   Pred. No. 24;
Matches  314;  Conservative    0;  Mismatches  355;  Indels  9;  Gaps  2;
```

```
Qy  107 GCGGCAAGGAGGGCCAACATCGCCCCGAACTGCCGCCCCCCCCGCAAGAGAAGGGGCTGCTGGA 166
            ||||| |||| |||||||||||| |||| ||  |||    ||||| |||  |||| |||
Db   99 GCGCGGCGGCGGCGCGGAGATGGCCGGCAGCCACCTCGACGAGGTGAAGCGCATGGTGGCGC 158

Qy  167 AGTGCGGCAAGGAGGAGGGCCAGCAGATGAAGGACTGCACGAGCGCCAACTTCTTCC 226
            || ||  || ||| ||||| |||| ||||||| || |||| ||||||| |||||
Db  159 AGTTCCGCGAGCCGCTGGTCAAGATCCAGGGCGGCCCGCCGCAGCCGCAGCGGCCG 218

Qy  227 GCGAGGACCTGCCCTTCCCCAGGGCAAGGCCCCGGCAGTTCCCCAGCGAGCGAGAACCGCG 286
            |||| ||||||||| || || ||| || ||||  ||  |||| ||||| || || ||
Db  219 CCGTCGCCCAGGCCAAGGACGCGCGCCCGGCGTCCGCGGCGCTCGAGCTGCAGCGAGGAGGGCCGCC 278

Qy  287 CCAAACAGCCCCACCAGCGGCGCGGCGCGGCTGGAGGTGCGCGGGCGGACAACCCCCCGGAGGAGGCCG 346
            ||||| ||||| |||| |||| ||| || || ||||| ||| | | |  |||
Db  279 CCCGGCGTCAAGGCGCCAGCAGCGAGTGGATCCTCAACTGCATCGCCCCACGGGCGGCGGCGACATCT 338

Qy  347 GCGGCCGAGCGGCGCCAGGCGCCACCCGCACCCTGACAACTTCCCCCAGATCACCCGGTGGCAGCGCCCCTGG 406
            ||||| | |||| ||| ||| |||||| |||  || ||  ||| | |||| |||
Db  339 ACGGCGTCACCACCGGCGTTCGGCGGCACCTCCGCGGCACCGGCGCCGA 398

Qy  407 TGAGCATCAAGTGGGCGGCGCAGATCAAGAGGAGGCCCTGCTGTGGACACCGGCCGGA---CG 463
            | || || |||| || ||| |||  |||  |||  ||| ||||||  ||||||||||    ||
Db  399 TCCAAGTGGAGCTCCTCAGGCATCTCAACGGCCGGAATCTTCGGCGACTGGCTCCGATGGCC 458

Qy  464 ACACCCGTGCTGCGAGGAGGATGAGCCTCGCCCGGCAAGTGGAAGCCCAAGATGATCGCGGCA 523
            |||| || ||| |||| ||| | || ||| ||  || ||  || ||||||||||| ||
Db  459 ACACGGCTGCGCGTCGGAGACGGTGCGGCGGGCGCCATGCCATCGCGTGCGCCATCAACACCCTTCCTCC 518

Qy  524 TCGGGCGGCTTCATCAAGGTGCGCCAGTAGCACCAGATCCTGATCGAGATCTCGCGGCAAGA 583
            |||| ||| ||||| || ||||| ||  |||| ||||| |||  |||| |||| ||| ||
Db  519 AGGGGCTACTCCGGCATCCGGTTGCGAGATCCTCGAGCCATCACCAAGAGCTGCTCAACACCG 578

Qy  584 AGGGCCATCGGCACCCGTGCTGATCGCTGATCGTCGATCATCGGCCGCCGCAACATGC 643
            ||||||| ||| ||| || ||| | || | || |||  |||  |||  | ||| |||
Db  579 GCGGTCACGCGGCCCTGCGCCGTCCCTGCGGCCACCATCACCGGTGACCTGGTTCCCC 638

Qy  644 TGACCCAGCTGGGCTGGCTGCACCCCTGAACTTCCCATGCCCCGAACTTCCCATCGCCGA 703
            || || || || |  ||||| ||  || || |||| || |||  ||| || ||| |||
Db  639 TGTCCTACATTGCCCGGCGCCTCATCATCACCGCGGCCCAACGCGCAGCGCCATCTCGCCCCGACG 698

Qy  704 AGCTGAAGCCCCGGCATGGACCGCCCAAGGTGAAGCAGTCGCCCTGACCGAGGAGAAGA 763
            |||||||||  |  ||| | |||    |||  |||  |||| ||| ||||| ||| |
Db  699 GCAGGAAG------GTGGACGACGCCGCGCGGGCATCGAGCGGCGTTNCAGCTGCGCCATCGAGCGGGCGG 752

Qy  764 TCAAGGCCCTGACCGCCA 781
            | ||||||| || ||
Db  753 TCTTCACGCTGAACCCCA 770
```

```
LOCUS       OGWCO51TH ZM_0.7_1.5_KB     753 bp    DNA    linear   GSS 19-JUN-2003
DEFINITION  OGWCO51TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMa0539J05,
            genomic survey sequence.
ACCESSION   CC675888
VERSION     CC675888.1  GI:32080584
KEYWORDS    GSS.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 753)
AUTHORS     Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohling,T.,
            Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
TITLE       Consortium for Maize Genomics
JOURNAL     Unpublished (2002)
COMMENT     Other GSSs: OGWCO51TV
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES            Location/Qualifiers
    source          1..753
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone="ZMMBMa0539J05"
                    /clone_lib="ZM_0.7_1.5_KB"
                    /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
                    methylation filtered genomic DNA library"
ORIGIN

Query Match       2.9%;    Score 72.4;   DB 29;   Length 753;
Best Local Similarity   46.2%;   Pred. No. 25;
Matches  281;  Conservative    0;  Mismatches  321;  Indels  6;  Gaps  1;
```

```
Qy   31 CAGGCCACCAGCGCCAACATCCTGATGCAGCCAGCAGCCGCAACTTCAAGGGCCCCAAGCGGATC 90
            ||||| ||| ||| || || || | || |||| ||  ||  |  | |||   ||||   |
Db   40 CACACCACCGACGACGACACCTTCGGCGGGTCCGGTCGTCAAGTGACCTCCACCAGCGCACC 99

Qy   91 ATCAAGTGCTTCAACTGGCCAAGGAGGGGCCAACATCGCCGGCAAGCTGCCGGCGGCCGGCGGC 150
            || ||| |||||||    || || |||| ||  |||  ||  ||  ||| || | |||  ||
Db  100 GTCGACCCGAGCCGCAGCGGGCGGCGGGCGTAACCGTACAACATATTCGGCGGGCGGCGGC 159

Qy  151 AAGAAGGGGCTGCTGGAAGTGCGGCAAGGAGGGGCAAGGAGGACTGCACCGAGCGC 210
            |  || ||| || |  |||| |||  ||||  | ||||  || |||| ||  ||||
Db  160 GACCAGCGGCGGGCCCTCGAGCTCGAGCTTCCCGGGCACCGGCAGCAGCGGGGACCTCGTCGTCCACCACCAC 219

Qy  211 CAGGCCAACTTCTTCCCGGAGACATCCTGATGCAGCCGCTTCCCGGAGAGCCCGAGTTCCCC 270
            ||||| ||| |  ||| || |  ||| | |||  |||| || || ||    || ||||
Db  220 TACATCCGCACCTCATCGACGAGGGCACCCCGATGCGGCTCAGTGCGGCTCAAGTGCGGCGCCGC 279

Qy  271 AGCGGAGCAGAACCGGCCAAACAGCCCCATCGCGGGGGCCCTGCCGGTGCGGCGGCGACAAC 330
            |||| ||| | |||   || |||||| | ||||  |||  ||||  ||  ||  || || |
Db  280 CTCTACACCCAACCGCGGCCTGACGACGACCACCGGCCGCGCCATCTCGCCCCGACG 339

Qy  331 CCCGCAGCGGAGGCGGCGGCGCCGAGCGCCCCCAAGGTGAAGGGGGCCCTGACCGAGGAGAAGA 390
            ||||  || | |||| ||||  | |||||   |||  | | |  |||||| ||||||  |||  |
Db  340 GCCTTCGGCGACGCGCCGGCGCGGGCAGCGGCGTTNCAGCTGCGCCATCGAGCGGGCGG 399

Qy  391 TGGCAGCGCCCCCTGGTGAGCATCAAGTGGGCGGCGGCCAGATCAAGAGGAGGCCCTGCTGGAC 450
            |||| || | ||| | ||   |||  |  ||| ||| |||  || |  | ||  ||| |||||
Db  400 GTCCGCGCGGCGACCCTGCTCGCCTTCGCCCGGCGCGGGGCGCCACTACGCGCGCCGTCGGGCCGC 459
```

```
Qy    451 ACCGGCGCCGACGACACCGTGCTGGAGGAGGATGAGCCTGCCCGGCAAGTGGAAGCCCAAG 510
              |---|--||||||    ||| ||||||||||||- | ||| |||| |||
Db    460 GCCTGGAAGCGGGGGGTACCTGCTCCACG-----GCCCGCCCGGCACGGGCAAGACCAGC 513

Qy    511 ATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCTGATCGAG 570
              ||||-|-|||||| ||-|-|    |-|||||-|||-||--| || |--|
Db    514 CTCGTCGCTGCCATCGCCAACCTGCTTCGACTTCGACGTCTACGACCTGGAGCTCACCACG 573

Qy    571 ATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGTGAACATCATC 630
              -|-|||  ||| |||| |||| |||-| ||||| ||| |||
Db    574 GTGCCCACCAACTCCCACCTCCCGCCGCTCGTCTCCACCACGCCCAGTCCGTCGTC 633

Qy    631 GGCCGCAA 638
              |-|-||
Db    634 GTCGTCGA 641
```

Search completed: April 10, 2004, 16:16:27
Job time : 4130.92 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    April 10, 2004, 02:45:46 ; Search time 619.487 Seconds
                                      (without alignments)
                                      16849.133 Million cell updates/sec

Title:           US-09-610-313-32
Perfect score:   2457
Sequence:        1 gtcgacgccaccatggccga.........gggctagccacggtgaattc 2457

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:   6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :  N_Genesec_29Jan04:*
            1: geneseqn1980s:*
            2: geneseqn1990s:*
            3: geneseqn2000s:*
            4: geneseqn2001as:*
            5: geneseqn2001bs:*
            6: geneseqn2002s:*
            7: geneseqn2003as:*
            8: geneseqn2003bs:*
            9: geneseqn2003cs:*
           10: geneseqn2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2457	100.0	2457	6	ABL39961	Abl39961 Synthetic
2	2443.4	99.4	2445	7	ACA03546	Aca03546 Synthetic
3	2443.4	99.4	2445	9	ADC13264	Adc13264 DNA of HI
4	2436.4	99.2	3930	9	ADC13230	Adc13230 DNA of HI
5	2436.2	99.2	2463	6	ABL39960	Abl39960 Synthetic
6	2434.8	99.1	3930	9	ADC13231	Adc13231 DNA of HI
7	2434.8	99.1	3930	9	ADC13232	Adc13232 DNA of HI
8	2434.8	99.1	5184	7	ACA03591	Aca03591 Synthetic
9	2434.8	99.1	5184	9	ADC13279	Adc13279 DNA of HI
10	2428.6	98.8	2457	7	ACA03547	Aca03547 Synthetic
11	2428.6	98.8	2457	9	ADC13265	Adc13265 DNA of HI
12	2415.4	98.3	2469	6	ABL39959	Abl39959 Synthetic
13	2404.4	97.9	3531	9	ADC13234	Adc13234 DNA of HI
14	2403.4	97.8	2457	7	ACA03548	Aca03548 Synthetic
15	2403.4	97.8	2457	9	ADC13266	Adc13266 DNA of HI
16	2402.8	97.8	3537	9	ADC13236	Adc13236 DNA of HI
17	2401.8	97.8	5145	7	ACA03521	Aca03521 Synthetic
18	2401.8	97.8	5145	9	ADC13233	Adc13233 DNA of HI
19	2391.8	97.3	3538	9	ADC13235	Adc13235 DNA of HI
20	2343.4	95.4	3607	7	ACA03551	Aca03551 Synthetic
21	2343.4	95.4	3607	9	ADC13269	Adc13269 DNA of HI
22	2325.2	94.6	3597	7	ACA03549	Aca03549 Synthetic
23	2325.2	94.6	3597	9	ADC13267	Adc13267 DNA of HI
24	2295.8	93.4	3624	7	ACA03550	Aca03550 Synthetic
25	2295.8	93.4	3624	9	ADC13268	Adc13268 DNA of HI
26	2136	86.9	2460	7	ACA03541	Aca03541 Synthetic
27	2136	86.9	2460	7	ACC78505	Acc78505 HIV p2Pol
28	2135	86.9	3564	7	ACC78488	Acc78488 HIV GagPo
29	2135	86.9	3564	7	ACC78489	Acc78489 HIV GagPo
30	2134.4	86.9	4716	7	ACA03522	Aca03522 Synthetic
31	2134.4	86.9	4716	9	ADC13238	Adc13238 DNA of HI
32	2130.6	86.7	3999	7	ACC78484	Acc78484 HIV GagCo
33	2129	86.7	3999	7	ACC78485	Acc78485 HIV GagCo
34	2129	86.7	3999	7	ACC78486	Acc78486 HIV GagCo
35	2129	86.7	5283	7	ACA03584	Aca03584 Synthetic
36	2129	86.7	5283	7	ACC78529	Acc78529 HIV TatRe
37	2128.8	86.6	4713	7	ACA03592	Aca03592 Synthetic
38	2128.8	86.6	4713	9	ADC13280	Adc13280 DNA of HI
39	2128.4	86.6	3462	9	ADC13237	Adc13237 DNA of HI
40	2115.2	86.1	2466	7	ACA03542	Aca03542 Synthetic
41	2115.2	86.1	2466	7	ACC78506	Acc78506 HIV p2Pol
42	2103	85.6	3735	7	ACA03545	Aca03545 Synthetic
43	2103	85.6	3735	9	ACC78509	Acc78509 HIV p2Pol
44	2096	85.3	2472	7	ACA03543	Aca03543 Synthetic
45	2096	85.3	2472	7	ACC78507	Acc78507 HIV p2Pol

ALIGNMENTS

```
RESULT 1
ABL39961
ID   ABL39961 standard; DNA; 2457 BP.
XX
AC   ABL39961;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Synthetic construct PR975YMWM SEQ ID NO:32.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO20204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J,  Barnett SW,  Engelbrecht S,  Van Rensburg EJ;
XX
DR   WPI; 2002-15420/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 10; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
```

```
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 U; 0 Other;

Query Match         100.0%;  Score 2457;  DB 6;  Length 2457;
Best Local Similarity  100.0%;  Pred. No. 6.9e-296;
Matches 2457;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Qy       1  GTCGACGCCACCATGGCCATGAGCCAGGCCAGGCCACCAGCCAACATCCTGATGCAG    60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  GTCGACGCCACCATGGCCATGAGCCAGGCCAGGCCACCAGCCAACATCCTGATGCAG    60

Qy      61  CGCAGCAACTTCAAGGGCCCCAAGCCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC   120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61  CGCAGCAACTTCAAGGGCCCCAAGCCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC   120

Qy     121  CACATCGCCCGCAACTGCCCGCGCCCCCCCGCGCCCCCCCGAGAAGAAGGGCTGCTGAAGGAG   180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  CACATCGCCCGCAACTGCCCGCGCCCCCCCGCGCCCCCCCGAGAAGAAGGGCTGCTGAAGGAG   180

Qy     181  GGCCACCAGATGAAGACTGCACCGGACGGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCC   240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  GGCCACCAGATGAAGACTGCACCGGACGGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCC   240

Qy     241  TTCCCCCAGGGCAAGGCCGGGCCGAGTTCCCCAGCGAGCAGCAACCGCGCGACACAGCCCCACC   300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  TTCCCCCAGGGCAAGGCCGGGCCGAGTTCCCCAGCGAGCAGCAACCGCGCGACACAGCCCCACC   300

Qy     301  AGCCGCGGAGCTGCAGGTGCGCGGCGGCGTGCGCAACCCCCGCAGCGAGGCCGGCGCCCAG   360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301  AGCCGCGGAGCTGCAGGTGCGCGGCGGCGTGCGCAACCCCCGCAGCGAGGCCGGCGCCCAG   360

Qy     361  GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCGCCCCCTGTGAGCATCAAGGTG   420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCGCCCCCTGTGAGCATCAAGGTG   420

Qy     421  GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCACCGGCGGCGGCCTGGACACCACCGGCGGCGGCCTG   480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421  GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCACCGGCGGCGGCCTGGACACCACCGGCGGCGGCCTG   480

Qy     481  ATGAGCCGCCGGGCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG   540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481  ATGAGCCGCCGGGCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG   540

Qy     541  GTGCGCGCCAGTACGACCAGATCGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG   600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541  GTGCGCGCCAGTACGACCAGATCGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG   600

Qy     601  CTGATCGGCCCCGTGAACATCATCGGTGAACATCTGCGACCTGACCCAGTCTGGGCTGC   660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601  CTGATCGGCCCCGTGAACATCATCGGTGAACATCTGCGACCTGACCCAGTCTGGGCTGC   660

Qy     661  ACCCTGAACTTCCCCATCAGCCCCATGAGCAGCCGTGAAGCTGAAGCTGAAGCTGAAGCTGAAG   720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661  ACCCTGAACTTCCCCATCAGCCCCATGAGCAGCCGTGAAGCTGAAGCTGAAGCTGAAGCTGAAG   720

Qy     721  GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCCTGACCGCC   780
            |||||||||||||||||||||||||||||||||||||||||||
Db     721  GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCCTGACCGCC   780

Qy     781  ATCTGCGAAGGATGGAAGAGGAGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC   840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781  ATCTGCGAAGGATGGAAGAGGAGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC   840

Qy     841  AACACCCCGTGTTCGCCATCAAGAGAGGACAGCACCAAGGTGGTGTGGTGCGAC   900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841  AACACCCCGTGTTCGCCATCAAGAGAGGACAGCACCAAGGTGGTGTGGTGCGAC   900

Qy     901  TTCCGCGAGCTGAACAAGCGCCACCAGACGCACCCCAGACTTCTGCGGAGGTGGCATCCCCAC   960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901  TTCCGCGAGCTGAACAAGCGCCACCAGACGCACCCCAGACTTCTGCGGAGGTGGCATCCCCAC   960
```

```
Qy     961  CCCGCCGCCTGAAGAAGAAGAAGAGAGCGTGACGTGGCGACGCCTACTTC  1020
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db     961  CCCGCCGCCTGAAGAAGAAGAAGAGAGCGTGACGTGGCGACGCCTACTTC  1020

Qy    1021  AGCGTGCCCCTGACGAGGACTTCCGCGAAGTACACCGCCTTCACCATCCCCAGCATCAAC  1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021  AGCGTGCCCCTGACGAGGACTTCCGCGAAGTACACCGCCTTCACCATCCCCAGCATCAAC  1080

Qy    1081  AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCAGGCTGAGAGGGCAGC  1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081  AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCAGGCTGAGAGGGCAGC  1140

Qy    1141  CCCAGCATCTTCAGAGCAGCATGACCAAGATCCTGGAGCGCCTTCCGCGCCCGCAACCCC  1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141  CCCAGCATCTTCAGAGCAGCATGACCAAGATCCTGGAGCGCCTTCCGCGCCCGCAACCCC  1200

Qy    1201  GAGATCGTGATCTACCAGGCCCCGTACGTGGGCAGCGACCTGGAGATCGGCCAGCGAC  1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201  GAGATCGTGATCTACCAGGCCCCGTACGTGGGCAGCGACCTGGAGATCGGCCAGCGAC  1260

Qy    1261  CGGCCAAGATGAGGAGCTGCGGCAAGCACCTGTGCTGCGGCTTCACCACCCCCGAC  1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261  CGGCCAAGATGAGGAGCTGCGGCAAGCACCTGTGCTGCGGCTTCACCACCCCCGAC  1320

Qy    1321  AAGAAGCACCAGAAGAGGAGCCCCCCTTCCTGCGCCATCGAGCTGCACCCCGACAAGTGACC  1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321  AAGAAGCACCAGAAGAGGAGCCCCCCTTCCTGCGCCATCGAGCTGCACCCCGACAAGTGACC  1380

Qy    1381  GTGCAGCCCCATCGAGCTGCCCGAGAAGCAGATCGGCCGTGACCGTGAACGACATCCAGAAGCTG  1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381  GTGCAGCCCCATCGAGCTGCCCGAGAAGCAGATCGGCCGTGACCGTGAACGACATCCAGAAGCTG  1440

Qy    1441  GTGGGCAAGTGAACTGGCCAGCGCAGATCTACCCCGGCATCGCGGCCAGCTGTGC  1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441  GTGGGCAAGTGAACTGGCCAGCGCAGATCTACCCCGGCATCGCGGCCAGCTGTGC  1500

Qy    1501  AAGCTGCTGCGCGGCGCAAGGGCCCTGACCGCATCGTGCCCCTGACCGTGACCGGAGGGCCAG  1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501  AAGCTGCTGCGCGGCGCAAGGGCCCTGACCGCATCGTGCCCCTGACCGTGACCGGAGGGCCAG  1560

Qy    1561  CTGGAGCTGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCTGCACGGCGTGTACTACGAC  1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561  CTGGAGCTGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCTGCACGGCGTGTACTACGAC  1620

Qy    1621  CCCAGCAAGGACCTGGTGGCAAGGCCAAGATCGGGCAAGAGACCGGACCTACCAG  1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621  CCCAGCAAGGACCTGGTGGCAAGGCCAAGATCGGGCAAGAGACCGGACCTACCAG  1680

Qy    1681  ATCTACCAGGAGCCCTTCAAGAACCTGAAGAACCTGAGAGACCTGCCAAGATGCGACCACCGCC  1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681  ATCTACCAGGAGCCCTTCAAGAACCTGAAGAACCTGAGAGACCTGCCAAGATGCGACCACCGCC  1740

Qy    1741  CACACCAACGACGTGAAGCAGTCAGCGTGACCGGCGTGCAGAAGATCGCCATGGAGGATC  1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741  CACACCAACGACGTGAAGCAGTCAGCGTGACCGGCGTGCAGAAGATCGCCATGGAGGATC  1800

Qy    1801  GTGATCTGGGCAAGCAGTGCCCCTGCGCCGCCATCAGAAGGAGACCTGGGAGACC  1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801  GTGATCTGGGCAAGCAGTGCCCCTGCGCCGCCATCAGAAGGAGACCTGGGAGACC  1860

Qy    1861  TGGTGGACGGACTACTGGCTGTGGTTACCAGCTGGAGGTTCGTGAACACCCCCC  1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861  TGGTGGACGGACTACTGGCTGTGGTTACCAGCTGGAGGTTCGTGAACACCCCCC  1920

Qy    1921  CCCCTGGTGAAGCTGTGGTGTACCAGGCGGAGAAGGCCCATCGCGGCGGCAAGACCTTC  1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921  CCCCTGGTGAAGCTGTGGTGTACCAGGCGGAGAAGGCCCATCGCGGCGGCAAGACCTTC  1980

Qy    1981  TACGTGGACGGCGCCGCCAACCGCGGCAAGGCCGCTACGTGACCGAC  2040
            |||||||||||||||||||||||||||||||||||||||||||||||
Db    1981  TACGTGGACGGCGCCGCCAACCGCGGCAAGGCCGCTACGTGACCGAC  2040
```

CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g., a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
XX
SQ   human immunodeficiency virus immunogenic peptide

Sequence 2445 BP; 562 A; 835 C; 751 G; 297 T; 0 U; 0 Other;

Query Match   99.4%;   Score 2443.4;   DB 7;   Length 2445;
Best Local Similarity   100.0%;   Pred. No. 3.3e-294;
Matches 2444;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
Qy         7 GCCACCATGGCCGAGGCCATGAGCGAGGCAGGCCAACATCCTGATGCAGCGCGAGC 66
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GCCACCATGGCCGAGGCCATGAGCGAGGCAGGCCAACATCCTGATGCAGCGCGAGC 60

Qy        67 AACTTCAAGGGCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 126
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 AACTTCAAGGGCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC 120

Qy       127 GCCCGCAACTGCCGCGCCCCCCGCCAAGAGGGCTGCTGAAGTGCGGCAAGGAGGGCCAC 186
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 GCCCGCAACTGCCGCGCCCCCCGCCAAGAGGGCTGCTGAAGTGCGGCAAGGAGGGCCAC 180

Qy       187 CAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCCGCGGAGGACTGGCCTTCCCC 246
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCCGCGGAGGACTGGCCTTCCCC 240

Qy       247 CAGGGCAAGGCCCCGGAGTTCCCAGCAGCAGAACCGCGCCAACAGCCCCACCAGCGCGC 306
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 CAGGGCAAGGCCCCGGAGTTCCCAGCAGCAGAACCGCGCCAACAGCCCCACCAGCGCGC 300

Qy       307 GAGCTGCAGTGCGGGCGACAACCCCCGCAGCGGAGGCCGCGCCGAGCGCCAGGGCACC 366
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GAGCTGCAGTGCGGGCGACAACCCCCGCAGCGGAGGCCGCGCCGAGCGCCAGGGCACC 360

Qy       367 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTGGCGGC 426
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTGGCGGC 420

Qy       427 CAGATCAAGGAGGACCCGTGCTGGGACACCCGTGCTGGAGGAGGAGATGAGC 486
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 CAGATCAAGGAGGACCCGTGCTGGGACACCCGTGCTGGAGGAGGAGATGAGC 480

Qy       487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGCTTCATCAAGGTGCGC 546
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGCTTCATCAAGGTGCGC 540

Qy       547 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC 606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 CAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATC 600

Qy       607 GGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG 666
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 GGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTG 660

Qy       667 AACTTCCCCATCAGCCCATCGAGACCTGCCCGGAGACCGTGCCCGTGAAGCTGAAGCGC 726
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AACTTCCCCATCAGCCCATCGAGACCTGCCCGGAGACCGTGCCCGTGAAGCTGAAGCGGC 720

Qy       727 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGGCCATCTGC 786
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 CCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGGCCATCTGC 780

Qy       787 GAGGAGATGGAGAAGGAGGGCCCGAAGATCGGCCCGGCAAGATCGGCCCGGAGAACCCTACAACACC 846
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 GAGGAGATGGAGAAGGAGGGCCCGAAGATCGGCCCGGCAAGATCGGCCCGGAGAACCCTACAACACC 840

Qy       847 CCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTTGGTGGACTTCCGC 906
```

Qy   2041 CGGGGCCGCAGAGATCGTGAGCCTGACCGAGCTGCGACCACCAACCAGAGAGACCGAGCTGCAG 2100
Db   2041 CGGGGCCGCAGAGATCGTGAGCCTGACCGAGCTGCGACCACCAACCAGAAGACCGAGCTGCAG 2100

Qy   2101 GCCATCCAGTGGCCCTGCAGGACAGCGGCAGCGGCGAGGTGAACATCGTGACCAGCCAG 2160
Db   2101 GCCATCCAGTGGCCCTGCAGGACAGCGGCAGCGGCGAGGTGAACATCGTGACCAGCCAG 2160

Qy   2161 TACGCCCTGGGCATCATCCAGGCCCAGCCCAGCCCACAAGAGCGAGCGAGCTGGTGAACCAG 2220
Db   2161 TACGCCCTGGGCATCATCCAGGCCCAGCCCAGCCCACAAGAGCGAGCGAGCTGGTGAACCAG 2220

Qy   2221 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCACAAG 2280
Db   2221 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCACAAG 2280

Qy   2281 GGCATCGGCGGCAACGAGCAGCAGCAGATCGACAAGCTGGTGAGCGCCATCCGCAAGGTGCTG 2340
Db   2281 GGCATCGGCGGCAACGAGCAGCAGCAGATCGACAAGCTGGTGAGCGCCATCCGCAAGGTGCTG 2340

Qy   2341 TTCCTGACGGGCATCGATGGCGGCATCGTGATCTACCAGTACACATGGACGACCTGTACGTG 2400
Db   2341 TTCCTGACGGGCATCGATGGCGGCATCGTGATCTACCAGTACACATGGACGACCTGTACGTG 2400

Qy   2401 GGCAGCGGCGGCGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2457
Db   2401 GGCAGCGGCGGCGCCCTAGGATCGATTAAAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2457

RESULT 2
ACA03546
ID   ACA03546 standard; DNA; 2445 BP.
XX
AC   ACA03546;
XX
DT   22-MAY-2003   (first entry)
XX
DE   Synthetic DNA ecoding immunogenic HIV peptide #29.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO200304657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 34; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for

```
Db  1921 GTGAAGCTGTGGTACCAGCTGGAGAGGAGGAGCCCATCATCGGCGCCGGCAACCGGAGACCTTCTACGTG 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  1987 GACGGGCGGCGCCAACCGCGAGGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGC 2046
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 GACGGGCGGCGCCAACCGCGAGGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGC 2040
Qy  2047 CGGCAAGAAGATCGTGAGCCTGACCGAGCACCAACCAGAGACCGAGCTGCAGCCGATC 2106
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 CGGCAGAAGATCGTGAGCCTGACCGAGCACCAACCAGAGACCGAGCTGCAGGCCATC 2100
Qy  2107 CAGCTGGCCCTGCAGGACAGCCGCAGCGAGTGAACATCGTGACCGACCAGTACGCC 2166
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CAGCTGGCCCTGCAGGACAGCCGCAGCGAGTGAACATCGTGACCGACCAGTACGCC 2160
Qy  2167 CTGGGCATCATCCAGGCCCGACAGCGAGAGCGAGAGCGAGAGCGTGGTGAACCAGATCATC 2226
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 CTGGGCATCATCCAGGCCCGACAGCGAGAGCGAGAGCGAGAGCGTGGTGAACCAGATCATC 2220
Qy  2227 GAGCAGTGATCAAGAAGGAGAAGGTGTACTGAGCTGGTGGGTGCCCGCCACAAGGGCATC 2286
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 GAGCAGTGATCAAGAAGGAGAAGGTGTACTGAGCTGGTGGGTGCCCGCCACAAGGGCATC 2280
Qy  2287 GGCGGCAACGAGCAGATCGACAAGGTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTG 2346
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 GGCGGCAACGAGCAGATCGACAAGGTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTG 2340
Qy  2347 GACGGCCATCGATGCGGCCATCGTGATCTGCCAGTACGATCATGGACGACCTGTACGTGGGCCAGC 2406
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GACGGCCATCGATGCGGCCATCGTGATCTGCCAGTACGATCATGGACGACCTGTACGTGGGCCAGC 2400
Qy  2407 GGCGGCCCTAGGATCGATTGATTAAAAAGCTTCCCGGGGCTAGCACCGGT 2451
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 GGCGGCCCTAGGATCGATTGATTAAAAAGCTTCCCGGGGCTAGCACCGGT 2445
```

RESULT 3
ADC13264
ID　ADC13264　standard; DNA; 2445 BP.
XX
AC　ADC13264;
XX
DT　18-DEC-2003　(first entry)
XX
DE　DNA of HIV construct p2Pol-opt-YMWM_C SEQ ID NO 43.
XX
KW　expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW　Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS　Human immunodeficiency virus.
XX
PN　WO2003004620-A2.
XX
PD　16-JAN-2003.
XX
PF　05-JUL-2002; 2002WO-US021420.
XX
PR　05-JUL-2001; 2001US-0303192P.
PR　31-AUG-2001; 2001US-0316860P.
PR　16-JAN-2002; 2002US-0349871P.
XX
PA　(CHIR) CHIRON CORP.
PA　(UYST-) UNIV STELLENBOSCH.
XX
PI　Zur Megede J,　Barnett SW,　Lian Y,　Engelbrecht S,　Van Rensburg EJ;
XX
DR　WPI; 2003-221593/21.
XX
PT　New expression cassette comprising a polynucleotide sequence encoding a
PT　polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT　Prot, or Rev polypeptide, useful for immunization, or generating
PT　packaging cell lines.
XX
PS　Disclosure; Fig 40; 301pp; English.

```
Db   841 CCCGGTGTTCGCCATCAAGAAGAAGAAGACACCAAGTGCGCAAGCTGGTGTGGACTTCCGC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   907 GAGCTGAACAAGGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 966
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAGCTGAACAAGGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 960
Qy   967 GGCCTGAAGAAGAAGAAGAAGACGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG 1026
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCTGAAGAAGAAGAAGAAGCGTGACCGTGCTGGACGTGGGCGACGCCTACTTCAGCGTG 1020
Qy  1027 CCCCTGGACAGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCGACATCAACAACGAG 1086
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CCCCTGGACCGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCGACATCAACAACGAG 1080
Qy  1087 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1146
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 ACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGC 1140
Qy  1147 ATCTTCCAGAGCAGCATGACCAAGATCCTCGAGCCCTTCCGCGCCCGCAACCCCGAGATC 1206
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 ATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCCGAGATC 1200
Qy  1207 GTGATCTACCAGGCCCCGTGTACGTGGGCAGCGACCTGGCAGCTGGGCAGCACGCGGCC 1266
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GTGATCTACCAGGCCCCGTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACGCGGCC 1260
Qy  1267 AAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCCCCGACACCCCGGAAG 1326
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 AAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCGACAACCCGGAAG 1320
Qy  1327 CACCAGAAGGAGCCCCCTTCCTGCCCATCGAGCTGCACCCCGACAGTGACCGTGCAG 1386
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CACCAGAAGGAGCCCCCTTCCTGCCCATCGAGCTGCACCCCGACAGTGACCGTGCAG 1380
Qy  1387 CCCATCGAGCTGCCCGAGAAGGAGAGCTGAACGACATCCAGAAGCTGGTGGGC 1446
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGC 1440
Qy  1447 AAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCCCAGCTGTGCCAAGCTG 1506
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 AAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCCGCAGCTGTGCCAAGCTG 1500
Qy  1507 CTGCGCGGCGGCCAAGGCCCTGCAAGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAG 1566
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CTGCGCGGCGGCCCGGCCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAG 1560
Qy  1567 CTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGC 1626
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 CTGGCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGC 1620
Qy  1627 AAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGACCTGGACCTGCCACCAGATCTAC 1686
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 AAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCAGGACCTGGACCTGCCACCAGATCTAC 1680
Qy  1687 CAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATGCGCACCGGCCCACACC 1746
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 CAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCCAAGATGCGCACCGGCCCACACC 1740
Qy  1747 AACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATC 1806
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 AACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATC 1800
Qy  1807 TGGGGCAAGACCCCCCCAAGTTCCGCGAGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTG 1866
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 TGGGGCAAGACCCCCCCAAGTTCCGCGAGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTG 1860
Qy  1867 ACCGACTACTGGCAGGCCACCTGAGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTG 1926
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 ACCGACTACTGGCAGGCCACCTGAGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTG 1920
Qy  1927 GTGAAGCTGTGGTACCAGCTGGAGAGGAGGAGCCCATCATCGGCGCCGGCAACCGGAGACCTTCTACGTG 1986
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

XX
CC The invention relates to a novel expression cassette comprising a
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC expression cassette can be used to treat HIV type C by gene therapy or
CC used in the development of a vaccine. The gene delivery vector is
CC administered intramuscularly, intramucosally, intranasally,
CC subcutaneously, intradermally, transdermally, intravaginally,
CC intrarectally, orally or intravenously. The expression cassette is useful
CC for immunisation, generating packaging cell lines and producing HIV
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC Type C related sequence of the invention.
XX
SQ Sequence 2445 BP; 562 A; 835 C; 751 G; 297 T; 0 U; 0 Other;

Query Match   99.4%;   Score 2443.4; DB 9; Length 2445;
Best Local Similarity 100.0%; Pred. No. 3.3e-294;
Matches 2444;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

```
PT   packaging cell lines.
XX
PS   Disclosure; Fig 6; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. the expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 890 A; 1365 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match         99.2%;  Score 2436.4;  DB 9;  Length 3930;
Best Local Similarity   100.0%;  Pred. No. 2.3e-293;
Matches 2437;  Conservative  0;  Mismatches  1;  Indels  1;  Gaps  0;

Qy       14 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCACCAGCGCCAGCATCCTGATGCAGCGGCAGCAACTTCA 73
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1487 TGGCCGAGGCCATGAGCCAGGCCACCAGCGCCACCAGCGCCAGCATCCTGATGCAGCGGCAGCAACTTCA 1546

Qy       74 AGGGGCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATGCCCCGGCA 133
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1547 AGGGGCCCAAGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCACATGCCCCGGCA 1606

Qy      134 ACTGCGGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGTGCGGCACCAGATGA 193
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1607 ACTGCGGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGTGCGGCACCAGATGA 1666

Qy      194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGGAGAACCGGCCAACAGCCCGCGCGAGCTGC 253
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1667 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGGGAGAACCGGCCAACAGCCCGCGCGAGCTGC 1726

Qy      254 AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGGCCAACAGCAGCCCGAGCGCCAGGCCGAGCTGC 313
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1727 AGGCCCGCGAGTTCCCCAGCGAGCAGAACCGGCCAACAGCAGCCCGAGCGCCAGGCCGAGCTGC 1786

Qy      314 AGGTGCGGCGGCGACAACCCCGCAGCGAGGCCGGCAAGAGGCCATCGGCGGCCACCCTGAACT 373
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1787 AGGTGCGGCGGCGACAACCCCGCAGCGAGGCCGGCAAGAGGCCATCGGCGGCCACCCTGAACT 1846

Qy      374 TCCCCCAGATCACCCTGTGCAGCGACATCAGGTGGGCGGCCAGATCA 433
            |||||||||||||||||||||||||||||||||||||||||||||||
Db     1847 TCCCCCAGATCACCCTGTGCAGCGACATCAGGTGGGCGGCCAGATCA 1906

Qy      434 AGGAGGCCCTGCTGGACACCGGCGACACCGGCGTGCTGGAGGATGAGCCTGCCCG 493
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1907 AGGAGGCCCTGCTGGACACCGGCGACACCGGCGTGCTGGAGGATGAGCCTGCCCG 1966

Qy      494 GCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAGTGCCGGCTACG 553
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAGTGCCGGCTACG 2026

Qy      554 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCATGGCCCCA 613
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2027 ACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCATGGCCCCA 2086

Qy      614 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2087 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 2146

Qy      674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 733
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2147 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 2206

Qy      734 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 793
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2207 TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 2266

Qy     1927 GTGAAGCTGTGTGTACCAGCTGGAGAGGAGCCATCGGCGCGCCGAGACCTTCTACGTG 1986
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GTGAAGCTGTGTGTACCAGCTGGAGAGGAGCCATCGGCGCGCCGAGACCTTCTACGTG 1980

Qy     1987 GACGGGCCGGCCAACCCGCGAGACCAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGC 2046
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 GACGGGCGCCGGCCAACCCGCGAGACCAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGC 2040

Qy     2047 CGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATC 2106
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 CGGCAGAAGATCGTGAGCCTGACCGAGACCAACCAGAAGACCGAGCTGCAGGCCATC 2100

Qy     2107 CAGTGCCCCTGCAGGACAGCGGCGCAGCGGAGGTGAACATCGTGACCGACAGCCAGTACGCC 2166
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 CAGTGCCCCTGCAGGACAGCGGCGCAGCGGAGGTGAACATCGTGACCGACAGCCAGTACGCC 2160

Qy     2167 CTGGGCATCATCCAGGCCCAGCCCGACAAGGAGGCCCAGCTGGAGCTGGTGAACCAGATCATC 2226
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 CTGGGCATCATCCAGGCCCAGCCCGACAAGGAGGCCCAGCTGGAGCTGGTGAACCAGATCATC 2220

Qy     2227 GAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATC 2286
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 GAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATC 2280

Qy     2287 GGCGGCAACGAGCAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTG 2346
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 GGCGGCAACGAGCAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTG 2340

Qy     2347 GACGGCATCGATGGCGGCATCGGCATCGTGATCTACCAGTACAATGGACGACCTGTACGTGGGCAGC 2406
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 GACGGCATCGATGGCGGCATCGGCATCGTGATCTACCAGTACAATGGACGACCTGTACGTGGGCAGC 2400

Qy     2407 GGCGGCCCTAGGATCGATTAAAAGCTTCCCCGGGGGTTAGCACCGGT 2451
            |||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 GGCGGCCCTAGGATCGATTAAAAGCTTCCCCGGGGGTTAGCACCGGT 2445


RESULT 4
ADC13230
ID   ADC13230 standard; DNA; 3930 BP.
XX
AC   ADC13230;
XX
DT   18-DEC-2003  (first entry)
XX
DE   DNA of HIV construct GagComplPolmut_C SEQ ID NO 9
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
```

```
Qy    794 TGGAGAAGGAGGGCAAGATCACCAAAGATCGGCCCCGAGAACCCTACAAACACCCCGTGT 853
          ||||||||||||||||||||||||| ||||||||||||||||||||||||| |||||||||
Db   2267 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACACACACCCCGTGT 2326

Qy    854 TCGCCATCAGAGAAGAGACAGCACCAAGTGGGCGACTGGTGGCACGTGGGACTTCCGCAGCGTGA 913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2327 TCGCCATCAGAGAAGAGACAGCACCAAGTGGGCGACTGGTGGCACGTGGGACTTCCGCAGCGTGA 2386

Qy    914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGCAGAGTGCAGTCGGGCCATCCCCACCCCGCGGGCCTGA 973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGCAGAGTGCAGTCGGGGCATCCCCACCCCGCGGGCCTGA 2446

Qy    974 AGAAGAAGAAGAGAGCGTGACCGCGTGACGTGGGCGACGCGCTCACTTCAGCGTGCCCCTGG 1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2447 AGAAGAAGAAGAGAGCGTGACCGCGTGACGTGGGCGACGCGCTCACTTCAGCGTGCCCCTGG 2506

Qy   1034 ACGAGGACTTCCGCAAGTACACCATCCCAGCATCACCATCCCAGCATCAACAACGAGACCCCCC 1093
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2507 ACGAGGACTTCCGCAAGTACACCATCCCAGCATCACCATCCCAGCATCAACAACGAGACCCCCC 2566

Qy   1094 GCATCCGTACCAGTACAAACGTGCTGCCCAGGGCTGAAGGCAGCCCAGCATCTTCC 1153
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2567 GCATCCGTACCAGTACAAACGTGCTGCCCAGGGCTGAAGGCAGCCCAGCATCTTCC 2626

Qy   1154 AGAGCAGCATGACCAAGATCCTGAGCGTTACTGGGCAGCGACCTGAGATGTGATCT 1213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2627 AGAGCAGCATGACCAAGATCCTGAGCGTTACTGGGCAGCGACCTGAGATGTGATCT 2686

Qy   1214 ACCAGGCCCCCCTGTACTGGGCAGCGACCACCTGCTCGCGTGGGGCTTCACCACCCCGAGATCG 1273
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2687 ACCAGGCCCCCCTGTACTGGGCAGCGACCACCTGCTCGCGTGGGGCTTCACCACCCCGAGATCG 2746

Qy   1274 AGGAGTGCTGCGCAAGCACACCTGCACCCCCGACAAGAAGCACCAGA 1333
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   2747 AGGAGTGCTGCGCAAGCACACCTGCACCCCCGACAAGAAGCACCAGA 2806

Qy   1334 AGGAGCCCCCCTTCTGCCCATCGAGCTGCACCCCGACAAGTGCAGCCCCATCG 1393
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2807 AGGAGCCCCCCTTCTGCCCATCGAGCTGCACCCCGACAAGTGCAGCCCCATCG 2866

Qy   1394 AGCTGCCCCAGAGGAGAGCTGAACCGTGAACACATCCAGAGAGCTGTGTGGGCAAGCTGA 1453
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2867 AGCTGCCCCAGAGGAGAGCTGAACCGTGAACACATCCAGAGAGCTGTGTGGGCAAGCTGA 2926

Qy   1454 ACTGGGCCCAGCCAGATCCTCACCCCCGACATCAAGTGCGCGTGACCGTGCTGCTCGGCG 1513
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2927 ACTGGGCCCAGCCAGATCCTCACCCCCGCATCAAGTGCGCGTGACCGTGCTGCTCGGCG 2986

Qy   1514 GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCCGAGCTGGAGCTGCGGCCG 1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2987 GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCCGAGCTGGAGCTGCGGCCG 3046

Qy   1574 AGAACCGCAGGATCCTGCGCGGAGCCCGTGACCGGCGTGTACTACGACACCCCAGCAAGGACC 1633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3047 AGAACCGCAGGATCCTGCGCGGAGCCCGTGACCGGCGTGTACTACGACACCCCAGCAAGGACC 3106

Qy   1634 TGGTGGCCGAGATCCAGAGCAGGCAGCAGGGCCACGACCTACCAGGAGC 1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3107 TGGTGGCCGAGATCCAGAGCAGGCAGCAGGGCCACGACCTACCAGGAGC 3166

Qy   1694 CCTTCAAGAACCTGAAGACCGCAAGTACGCCAAGATGCCAAGACCACCGCCCAACGACG 1753
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3167 CCTTCAAGAACCTGAAGACCGCAAGTACGCCAAGATGCCAAGACCACCGCCCAACGACG 3226

Qy   1754 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGGCCATGGAGCAGCATCGTGATCGGGGCA 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3227 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGGCCATGGAGCAGCATCGTGATCGGGGCA 3286

Qy   1814 AGACCCCCAAGTTCCGCCTGCCCATCCAGAGAGGAGAGACCTGGGAGACCTGTGGGGACACT 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3287 AGACCCCCAAGTTCCGCCTGCCCATCCAGAGAGGAGAGACCTGGGAGACCTGTGGGGACACT 3346
```

```
Qy   1874 ACTGGCAGGCCACCTGGATCCCCAGTGGGAGTTCGTGAACACCCCCCTGTGAAGC 1933
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3347 ACTGGCAGGCCACCTGGATCCCCAGTGGGAGTTCGTGAACACCCCCCCTGTGAAGC 3406

Qy   1934 TGTGGTACCAGTCGAGAAGAGAGCCCATCATCGGGCGCCGAGACCTTCTACGTGGACGCGCG 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3407 TGTGGTACCAGTCGAGAAGAGAGCCCATCATCGGGCGCCGAGACCTTCTACGTGGACGCGGCG 3466

Qy   1994 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGGGGCCCGGCAGA 2053
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3467 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGGGCCCGGCAGA 3526

Qy   2054 AGATCGTGAGCGTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 2113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3527 AGATCGTGAGCGTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 3586

Qy   2114 CCCTGCAGGACACGCGGCGACGGAGGTGAACATCGTGACCGCAGCAGCCAGTAGCCCTGGCA 2173
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3587 CCCTGCAGGACAGCGGCGACGGAGGTGAACATCGTGACCGCAGCCAGTAGCCCTGGGCA 3646

Qy   2174 TCATCCAGGCCCCAGCCCGACAAGAGCGAGAGCGAGACCTGGTGAACCAGATCATCGAGCAGC 2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3647 TCATCCAGGCCCCAGCCCGACAAGAGCGAGAGCGAGACCTGGTGAACCAGATCATCGAGCAGC 3706

Qy   2234 TGATCAAGAAGGAGAAGGTGTACTGGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGGCA 2293
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3707 TGATCAAGAAGGAGAAGGTGTACTGGGCCTGGGTGCCCGCCCACAAGGGCATCGGCGGGCA 3766

Qy   2294 ACGAGCAGATCGACAAGCTGGTGAGCGAGCAAGGGCATCCCAAGGTGCGTGTTCCTGGACGGCA 2353
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3767 ACGAGCAGATCGACAAGCTGGTGAGCGAGCAAGGGCATCCCAAGGTGCGTGTTCCTGGACGGCA 3826

Qy   2354 TCGATGGCGGATCGTGATCTACCAGTACATGGACCACCTGTACGTGGGCAGCGGCGGGCC 2413
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3827 TCGATGGCGGATCGTGATCTACCAGTACATGGACCACCTGTACGTGGGCAGCGGCGGGCC 3886

Qy   2414 CTAGGATCGATTAAAAGCTTCCGGGGGCTAGGACCCGGT 2451
          |||||||||||||||||||||||||||||||||||||||
Db   3887 CTAGGATCGATTAAAAGCTTCCGGGGGCTAGGACCACCGGT 3924
```

```
RESULT 5
ABL39960
ID   ABL39960 standard; DNA; 2463 BP.
XX
AC   ABL39960;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic construct PR975YM SEQ ID NO:31.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO20020493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.
XX
```

PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunisation or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 9; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 U; 0 Other;

Query Match   99.2%;   Score 2436.2;   DB 6;   Length 2463;
Best Local Similarity   99.6%;   Pred. No. 2.6e-293;
Matches 2454;   Conservative 0;   Mismatches   3;   Indels   6;   Gaps   1;

```
Qy      1 GTCGACGCCACCATGGCCGAGGCCATGAGCGCCAGCGCACCAGCCCAACATCCTGATGCAG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCGACGCCACCATGGCCGAGGCCATGAGCGCCAGCGCACCAGCCCAACATCCTGATGCAG 60

Qy     61 CGCAGCAACTTCAAGGGCCCAACTGCCCGGCCCCCCCCGCAAGAAGGCGTGCTGGAAGTGC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCAGCAACTTCAAGGGCCCAACTGCCCGGCCCCCCCCGCAAGAAGGCGTGCTGGAAGTGC 120
```

```
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 7; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 889 A; 1365 C; 1214 G; 462 T; 0 U; 0 Other;

Query Match    99.1%; Score 2434.8; DB 9; Length 3930;
Best Local Similarity  99.9%; Pred. No. 3.6e-293;
Matches 2436; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy      14 TGGCCGAGGCCATGAGCCACCAGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1487 TCGCCCGAGGCCATGAGCCACCAGCCACCAGCGCCAACATCCTGATGCAGCGCAGCAACTTCA 1546

Qy      74 AGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA 133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1547 AGGGCCCCAAGCCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCCCA 1606

Qy     134 ACTGCGGCCGCCCCCCGCCAAGAAGGGCGTCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1607 ACTGCGGCCGCCCCCCGCCAAGAAGGGCGTCTGGAAGTGCGGCAAGGAGGGGCCACCAGATGA 1666

Qy     194 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1667 AGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGGGCA 1726

Qy     254 AGGCCCGGAGTTCCCAGCGAGCAGAGAACCGGCCCAACAGCCCCACCAGCCCGCGAGCTGC 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1727 AGGCCCGGAGTTCCCAGCGAGCAGAGAACCGGCCCAACAGCCCCACCAGCCCGCGAGCTGC 1786

Qy     314 AGTGCGGCGGCACAACCCCGCAGCGAGGCCCGCGGCCGCGGCCCGAGCGCCACCCTGAACT 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1787 AGTGCGGCGGCACAACCCCGCAGCGAGGCCCGCGGCCGCGGCCCGAGCGCCACCCTGAACT 1846

Qy     374 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGTGAGCATCAAGGTGGCGGCCCAGATCA 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1847 TCCCCCAGATCACCCTGTGGCAGCGCCCCCTGTGAGCATCAAGGTGGCGGCCCAGATCA 1906

Qy     434 AGGAGGCCCTGCTGACACCGGCCGCCGCCAAGAAGCCATCGGCACCGTGCTGATCGGCCCCG 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1907 AGGAGGCCCTGCTGACACCGGCCGCCGCCAAGAAGCCATCGGCACCGTGCTGATCGGCCCCG 1966

Qy     494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCAGTACG 553
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCAGTACG 2026

Qy     554 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGCCATCGGCACCGTGCTGGCCCCA 613
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2027 ACCAGATCCTGATCGAGATCTGCGGCCAAGAAGCCATCGGCACCGTGCTGGCCCCA 2086
```

```
Db    1801 AGCATCCTGATCTGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAGGAGGACCTGG 1860
Qy    1855 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGGAGTTCGTGAAC 1914
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GAGACCTGGTGGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGGAGTTCGTGAAC 1920
Qy    1915 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGGCCCAG 1974
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGGCCCAG 1980
Qy    1975 ACCTTCTACGTGGACGGCGCCGCCAACCGCCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2034
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 ACCTTCTACGTGGACGGCGCCGCCAACCGCCGAGACCAAGATCGGCAAGGCCGGCTACGTG 2040
Qy    2035 ACCGAACCGGGGCCGGCCAGAGAGATCGTGAGCCTGAGCGACCACCAACCAGAGACCGAG 2094
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 ACCGAACCGGGGCCGGCCAGAGAGATCGTGAGCCTGAGCGACCACCAACCAGAGACCGAG 2100
Qy    2095 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGGAGGTGAACATCGTGACCGAC 2154
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 CTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGGAGGTGAACATCGTGACCGAC 2160
Qy    2155 AGCCAGTACGCCCTGGCCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGGAGCTGGTG 2214
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 AGCCAGTACGCCCTGGCCATCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGGAGCTGGTG 2220
Qy    2215 AACCAGATCATCGAGCAGTGATCAAGAAGGAGGAAGGGTGTACCTGAGCTGGTGCCGCGCC 2274
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGGAAGGGTGTACCTGAGCTGGTGCCCGCC 2280
Qy    2275 CACAAGGGCATCGGGCGGCAACGAGCAGATCGACAACTCGTGAGCAAGGGCATCCGCAAG 2334
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 CACAAGGGCATCGGGCGGCAACGAGCAGATCGACAACTCGTGAGCAAGGGCATCCGCAAG 2340
Qy    2335 GTGCTGTTCCTGGACAACGGCGCATCGATGGCGGCATCGTGATCTGATCGTGAGCCAGACCTG 2394
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GTGCTGTTCCTGGACGACGGCGCATCGATGGCGGCATCGTGATCTGATCGTGAGCCAGACCTG 2400
Qy    2395 TACGTGGGCAGCGCGGCGGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2454
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 TACGTGGGCAGCGCGGCGGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460
Qy    2455 TTC 2457
           |||
Db    2461 TTC 2463
```

```
RESULT 6
ADC13231
ID   ADC13231 standard; DNA; 3930 BP.
XX
AC   ADC13231;
XX
DT   18-DEC-2003  (first entry)
XX
DE   DNA of HIV construct GagComplPolmutAtt_C SEQ ID NO 10.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
```

```
Db  2087 CCCCCTGAACATCATCGGCGCCAACATGCTGACCGGATGGGGCTGCACCCTGAACTTCC 2146
Qy   674 CCATCAGCCCCATCGAGCCCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 733
Db  2147 CCATCAGCCCCATCGAGCGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 2206

Qy   734 TGAAGCAGTGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 793
Db  2207 TGAAGCAGTGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGCGAGGAGA 2266

Qy   794 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 853
Db  2267 TGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 2326

Qy   854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGACCTG 913
Db  2327 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGACCTGA 2386

Qy   914 ACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 973
Db  2387 ACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 2446

Qy   974 AGAAGAAGAAGAGCGTGCTGACCGTGCTGGGCGACGGCCTACTTCAGCGTGCCCCTGG 1033
Db  2447 AGAAGAAGAAGAGCGTGCTGACCGTGCTGGGCGACGGCCTACTTCAGCGTGCCCCTGG 2506

Qy  1034 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 1093
Db  2507 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 2566

Qy  1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCC 1153
Db  2567 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCGCCATCTTCC 2626

Qy  1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAATCGTGATCT 1213
Db  2627 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAATCGTGATCT 2686

Qy  1214 ACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACGCGCCCAAGATCG 1273
Db  2687 ACCAGGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCACGCGCCCAAGATCG 2746

Qy  1274 AGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 1333
Db  2747 AGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 2806

Qy  1334 AGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCG 1393
Db  2807 AGGAGCCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCG 2866

Qy  1394 AGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 1453
Db  2867 AGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 2926

Qy  1454 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCCAGCTGTGTGCGGCGCTGCGCG 1513
Db  2927 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCCAGCTGTGTGCGGCGCTGCGCG 2986

Qy  1514 GCGCCAAGGCCCTGACCGTGCCCCTGCGCGAGGAGGCCGAGCTGGAGCTGGCGCG 1573
Db  2987 GCGCCAAGGCCCTGACCGTGCCCCTGCGCGAGGAGGCCGAGCTGGAGCTGGCGCCG 3046

Qy  1574 AGAACCGCGGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACC 1633
Db  3047 AGAACCGCGGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACC 3106

Qy  1634 TGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGC 1693
Db  3107 TGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGC 3166

Qy  1694 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGGATGCGCGGCGCCCACACCAACGACG 1753
Db  3167 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCCGGATGCGCGGCGCCCACACCAACGACG 3226
```

```
Qy  1754 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCATGGAGAGCATCGTGATCTGGGGCA 1813
Db  3227 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCA 3286

Qy  1814 AGACCCCCAAGTTCCGGCTGCCTGCCATCCAGAAGGAGACCTGGGACCTGGTGGACCGACT 1873
Db  3287 AGACCCCCAAGTTCCGGCTGCCTGCCATCCAGAAGGAGACCTGGGACCTGGTGGACCGACT 3346

Qy  1874 ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGGTGAAGC 1933
Db  3347 ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCTGGTGAAGC 3406

Qy  1934 TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGCGCCGAGACCTTCTACGTGGACGGCG 1993
Db  3407 TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCG 3466

Qy  1994 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGCCGGCAGA 2053
Db  3467 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGCCGGCAGA 3526

Qy  2054 AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 2113
Db  3527 AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 3586

Qy  2114 CCCTGCAGGACAGCGGCCAGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 2173
Db  3587 CCCTGCAGGACAGCGGCCAGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 3646

Qy  2174 TCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGGTCGTGAACCAGATCATCGAGCAGC 2233
Db  3647 TCATCCAGGCCCAGCCCGACAAGAGCGAGAGCGAGGTCGTGAACCAGATCATCGAGCAGC 3706

Qy  2234 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 2293
Db  3707 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 3766

Qy  2294 ACGAGCAGATCGACAAGCTGGTGAGCAAGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 2353
Db  3767 ACGAGCAGATCGACAAGCTGGTGAGCAAGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 3826

Qy  2354 TCGATGGCGGCATCGTGATCTACCAGTACGATGACGACCTGTACGTGGGCAGCCGGCGGCC 2413
Db  3827 TCGATGGCGGCATCGTGATCTACCAGTACGATGACGACCTGTACGTGGGCAGCCGGCGGCC 3886

Qy  2414 CTAGGATCGATTAAAAGCTTCCCGGGGGGTAGCACCGGT 2451
Db  3887 CTAGGATCGATTAAAAGCTTCCCGGGGGGTAGCACCGGT 3924
```

PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR-) CHIRON CORP.
PA   (UVST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 8; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Prot, Tat, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3930 BP; 889 A; 1366 C; 1214 G; 461 T; 0 U; 0 Other;

Query Match   99.1%; Score 2434.8; DB 9; Length 3930;
Best Local Similarity   99.9%; Pred. No. 3.6e-293;
Matches 2436; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy    14 TGGCGGAGGCCATGAGCCAGGCCACCGGCGCCAACATCCTGATCGCCGACATCCTGATCGCCGACACTTCA 73
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1487 TGGCGGAGGCCATGAGCCAGGCCACCGGCGCCAACATCCTGATCGCCGACATCCTGATCGCCGACACTTCA 1546

Qy    74 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATGCCCGCA 133
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db  1547 AGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATGCCCCGA 1606

Qy   134 ACTGCGGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 193
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1607 ACTGCGGCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 1666

Qy   194 AGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGCGGAGGACCTGCCCTTCCCCCAGGGCA 253
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1667 AGGACTGCACCGAGCGGCCAGGCCAACTTCTTCCGCGGAGGACCTGCCCTTCCCCCAGGGCA 1726

Qy   254 AGGCCCGGAGTTCCCCAGCGGAGCAGAACCGCGCCAACAGCGGCCCGAGGGGCCGGAGCTGC 313
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1727 AGGCCCGGAGTTCCCCAGCGGAGCAGAACCGCGCCAACAGCGGCCCGAGGGGCCGGAGCTGC 1786

Qy   314 AGGTGCCGGCGGACACCCCGGCGGCCCCCTGGTGAGCGGGGCATCAAGGTGGGGGGCCAGACT 373
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1787 AGGTGCCGGCGGACACCCCGGCGGCCCCCTGGTGAGCGGGGCATCAAGGTGGGGGGCCAGACT 1846

Qy   374 TCCCCCAGATCACCCTGTGGCAGCGGCCCCTGGTGAGCGGGGCATCAAGGTGGGGGGCCAGATCA 433
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1847 TCCCCCAGATCACCCTGTGGCAGCGGCCCCTGGTGAGCGGGGCATCAAGGTGGGGGGCCAGATCA 1906

Qy   434 AGGAGGCCCTGCTGACACCGGCCGACACCGGCGCCGACACCGGC 493
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1907 AGGAGGCCCTGCTGACACCGGCCGACACCGGCGCCGACACCGGC 1956

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1967 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 2026

Qy   554 ACCAGATCCTGATCGAGATTGCGGCAAGAAGGCGATCGGCACCGTGCTGGTGGGCCCAACATCCCCG 613
```

```
Qy   614 CCCCCGTGAACATCATCGGCCGCAACATGTGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 673
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2087 CCCCCGTGAACATCATCGGCCAACATGTGCTGACCCAGCTGGGCTGCACCCTGAACTTCC 2146

Qy   674 CCATCAGCCCATGAGCCGTGAGCTGAAGCTGCCGGCATGGACGCGCCCAAGG 733
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2147 CCATCAGCCCATGAGCCGTGAGCTGAAGCTGCCGGCATGGACGCGCCCAAGG 2206

Qy   734 TGAAGAGTGGCCCCTGACCGAGGAGAATCAAGGCCCTGACCGCCATCTGCGAGGAGA 793
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2207 TGAAGCAGTGGCCCCTGACCGAGGAGAATCAAGGGCCTGACCGCCATCTGCGAGGAGA 2266

Qy   794 TGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 853
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2267 TGGAGGAGGGCAAGATCAGCAAGATCGGCCCCGAGAACCCCTACAACACCCCCGTGT 2326

Qy   854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGTCGGTGGACTTCCGCGAGCTGA 913
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2327 TCGCCATCAAGAAGAAGGACAGCAGCACCAAGTGGCGCAAGTCGGTGGACTTCCGCGAGCTGA 2386

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 973
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2387 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCGGCCTGA 2446

Qy   974 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCGGCCTACTTCAGCGTGCCCCTGG 1033
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2447 AGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCGGCCTACTTCAGCGTGCCCCTGG 2506

Qy  1034 ACGAGGACTTCCCCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 1093
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2507 ACGAGGACTTCCCCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 2566

Qy  1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCTGGAAGGGCGAACCCCAGGCATCTTCC 1153
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2567 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCTGGAAGGGCGAACCCCAGGCATCTTCC 2626

Qy  1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 1213
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2627 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 2686

Qy  1214 ACCAGGCCCCCGTGTACGTGGGCAAGGAGTGCGGCAAGGAGCGGCCAGCACCGGCAAGATCG 1273
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2687 ACCAGGCCCCCGTGTACGTGGGCAAGGAGTGCGGCAAGGAGCGGCCAGCACCGGCAAGATCG 2746

Qy  1274 AGGAGCTGCGCGCCGCCCACCTGCTGCGCGGCTTCACCCGCTGCGCCATCGAGCTGCAGAGAGCACCAGA 1333
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2747 AGGAGCTGCGCGCCGCCCACCTGCTGCGCGGCTTCACCCGCTGCGCCATCGAGCTGCAGAGAGCACCAGA 2806

Qy  1334 AGGAGCCCCCTTCCTGCTGCCCATCGAGTGCGACAAGAGCCGTGGACCGGCGCAGCTGC 1393
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2807 AGGAGCCCCCTTCCTGCTGCCCATCGAGTGCGACAAGAGCCGTGGACCGGCGCAGCTGC 2866

Qy  1394 AGTGCCCGAGAAGGAGAGCTGACCGACCGCATCAAGGTGGGGCAAGAGTCTGTGTGTGGGCAAGTGA 1453
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2867 AGTGCCCGAGAAGGAGAGCTGACCGACCGCATCAAGGTGGGGCAAGAGTCTGTGTGTGGGCAAGTGA 2926

Qy  1454 ACTGGGCCAGCCAGAAGGAGAGCTCACCCCGGCATCAAGGTGCGCGCGCCGAGCATCAAGGTGCCGAGATCA 1513
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2927 ACTGGGCCAGCCAGAAGGAGAGCTCACCCCGGCATCAAGGTGCGCGCGCCGAGCATCAAGGTGCCGAGATCA 2986

Qy  1514 GGCCCAAGGCCCTGACCGAGGTGGTGCCGCTGACCGAGGAGGGCGAGTGGCCTGCCCG 1573
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2987 GGCCCAAGGCCCTGACCGAGGTGGTGCCGCTGACCGAGGAGGGCGAGTGGCCTGGCCG 3046

Qy  1574 AGAACCGGGAGATCCTGCGCCAAGATAACGACCCCTGTACTACGACCCCAGCAAGGACC 1633
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3047 AGAACCGGGAGATCCTGCGCCAAGATGATCGGCTGTACTACGACCCCAGCAAGGACC 3106

Qy  1634 TGTTGGCCGAGATCGAAGAGCCAGAGCAGGGCCAGAGTGCGACCTACCAGGAGC 1693
```

PF 05-JUL-2002; 2002WO-US021421.
XX
PR 05-JUL-2001; 2001US-0303192P.
PR 31-AUG-2001; 2001US-0316860P.
PR 16-JAN-2002; 2002US-0349728P.
PR 16-JAN-2002; 2002US-0349793P.
PR 16-JAN-2002; 2002US-0349871P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Zur Megede J, Barnett SW, Lian Y;
XX
DR WPI; 2003-221602/21.
XX
PT New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT polypeptides, useful as immunogenic compositions or vaccines for
PT generating humoral or cellular immune responses against HIV in a subject,
PT especially humans.
XX
PS Example 1; Fig 79; 262pp; English.
XX
CC The invention describes a synthetic polynucleotide encoding 2 or more
CC immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC derived from different HIV subtypes. The polynucleotide is useful for
CC immunisation, generation of packaging cell lines, or production of HIV
CC polypeptides. The polynucleotide and its encoded proteins are useful as
CC immunogenic compositions or vaccines for generating humoral or cellular
CC immune responses against HIV in a subject, or for inducing neutralising
CC antibodies against HIV. The gene delivery vector comprising the
CC polynucleotide is also useful for DNA immunisation of, or for generating
CC an immune response (e.g. a humoral or cellular immune response) in, a
CC subject such as a mammal, particularly a human. This sequence encodes a
CC human immunodeficiency virus immunogenic peptide
XX
SQ Sequence 5184 BP; 1139 A; 1852 C; 1610 G; 583 T; 0 U; 0 Other;

```
Query Match             99.1%;  Score 2434.8;  DB 7;  Length 5184;
Best Local Similarity   99.9%;  Pred. No. 3.4e-293;
Matches  2436;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     14 TGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATCGCAGCGCAACTTCA   73
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 TGGCCGAGGCCATGAGCGCCAGGCCACCAGCGCCAACATCCTGATCGCAGCGCAACTTCA 2800

Qy     74 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA  133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCA 2860

Qy    134 ACTGCGCGCCCCCCGCAAGAGAGGGTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA  193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 ACTGCGCGCCCCCCGCAAGAGAGGGTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGA 2920

Qy    194 AGGACTGCACCGAGGCGCAGGCCAGCCACTTCTTCCGGCAGGAACCTGGCCTTCCCCAGGGGCA  253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGACTGCACCGAGGCGCAGGCCAGCCACTTCTTCCGGCAGGAACCTGGCCTTCCCCAGGGGCA 2980

Qy    254 AGGACCCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCGCCAACCAGCCCGAGCTGC  313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 AGGACCCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCGCCAACCAGCCCGAGCTGC 3040

Qy    314 AGGTGCGGCGGCGACAACCCCCGCAGCCCGAGGCGCCCAGCGCCCAGGCGCCAGCCCTGAACT  373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGGCGGCGACAACCCCCGCAGCCCGAGGCGCCCAGCGCCCAGGCGCCAGCCCTGAACT 3100

Qy    374 TCCCCAGATCACCCTGTGGCAGCGGCCATCAAGTGGGGCGGCATCAAGTGGGGCGGCATCA  433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCAGATCACCCTGTGGCAGCGGCCATCAAGTGGGGCGGCATCAAGTGGGGCGGCATCA 3160

Qy    434 AGGAGGCCCTGCTGGACACCGGCGACACCGGCTGCTGGAGGATGAGCGCCTGCCCG  493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGCCCTGCTGGACACCGGCACCGGCTGCTGGAGGATGAGCGCCTGCCCG 3220
```

```
Qy   1694 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCGCCACCAACCACG 1753
          |||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db   3167 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCGCCACCAACCACG 3226

Qy   1754 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCA 1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3227 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCA 3286

Qy   1814 AGACCCCCAAGTTCGCGCCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACT 1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3287 AGACCCCCAAGTTCCCGCCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACT 3346

Qy   1874 ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCCCCCTGGTGAAGC 1933
          |||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   3347 ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCCCCCTGGTGAAGC 3406

Qy   1934 TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGAGACCTTCTACGTGGACGGCG 1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3407 TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGAGACCTTCTACGTGGACGGCG 3466

Qy   1994 CGGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACGAGGCCGCGGCCAGA 2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3467 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACGAGGCCGCGGCCAGA 3526

Qy   2054 AGATCGTGAGCCTGACCGAGACCACCAACCAGCAGACCGAGTGCAGGCCATCATCCAGCTGG 2113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3527 AGATCGTGAGCCTGACCGAGACCACCAACCAGCAGACCGAGCTGCAGGCCATCATCCAGCTGG 3586

Qy   2114 CCCTGCAGGACAGCGGCGAGTGACATCGTGACCAGCGCCAGTACGCCCTGGCA 2173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3587 CCCTGCAGGACAGCGGCAGCGAGTGACATCGTGACCAGCGCCAGTACGCCCTGGCA 3646

Qy   2174 TCATCCAGGCCCCAGCCCGACAAGAGCCGAGCGGAGCGAGCTGGTGAACCAGATCATCGAGCAGC 2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3647 TCATCCAGGCCCCAGCCCGACAAGAGCCGAGCGGAGCGAGCTGGTGAACCAGATCATCGAGCAGC 3706

Qy   2234 TGATCAAGAAGGAGAAAGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCCGCA 2293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3707 TGATCAAGAAGGAGAAAGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 3766

Qy   2294 ACGAGCAGATCGACAAGCTGGTGAGCGCCGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 2353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3767 ACGAGCAGATCGACAAGCTGGTGAGCGCCGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 3826

Qy   2354 TCGATGGCCGGCATCGTCGATCATCACCAGTACGACATGGCAGCGGCGGCC 2413
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3827 TCGATGGCCGGCATCGTCGATCATCACCAGTACGACATGGCAGCGGCGGCC 3886

Qy   2414 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
          ||||||||||||||||||||||||||||||||||||||
Db   3887 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 3924
```

Db 3221 GCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGCTTCATCAAGGTGCGCCAGTACG 3280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   554 ACCAGATCCTGATCGAGATCGCGGCCAAGAGGCCACGAGCGGCCAGTACCGTGCGCCCCA 613
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3281 ACCAGATCCTGATCGAGATCTCGCGCGAAGAGGCCATCGGCCACCGTGCTGCTGGCGCCCA 3340

Qy   614 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGTGGGCTGACCCTGAACTTCC 673
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3341 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGTGGGCTGACCCTGAACTTCC 3400

Qy   674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGTCAGCCCGGCACGGCCCCAAGG 733
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3401 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGTCAGCCCGGCACGGCCCCAAG 3460

Qy   734 TGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGCGAGAGA 793
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3461 TGAAGCAGTGGCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGCGAGAGA 3520

Qy   794 TGGAGAAGGAGGGCAAGATCACCAAGGAGATCGGCCCCGAGAACCCTACACACCCCCGTGT 853
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3521 TGGAGAAGGAGGGCAAGATCACCAAGGAGATCGGCCCCGAGAACCCTACACACACCCCGTGT 3580

Qy   854 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 913
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3581 TCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGA 3640

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCGCCGCCGGCCTGA 973
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3641 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCGCCGCCGGCCTGA 3700

Qy   974 AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGCGAGCAAGTGGCCTACTTCAGCGTCCCCTGG 1033
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3701 AGAAGAAGAGAGCGTGACCGTGCTGGACGTGGCGAGCAAGTGGCCTACTTCAGCGTCCCCTGG 3760

Qy 1034 ACGAGGACTTCGCCAAGTACACCGCCTTCACCACCATCCCCAGCAATCAACAACGAGACCCCCG 1093
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3761 ACGAGGACTTCGCCAAGTACACCGCCTTCACCACCATCCCCAGCATCAACAACGAGACCCCCG 3820

Qy 1094 GCATCCGTACCAGTACACACGTTGCTGCCCCGGCTGGAAGGGCAGCCCAGCCATCTTCC 1153
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3821 GCATCCGTACCAGTACACACGTGCTGCCCCGGCTGGAAGGGCAGCCCAGCCATCTTCC 3880

Qy 1154 AGAGCAGCATGACCAAGATACCTCGAGCCCTTCCGCGCCGCAACCCCGCCAGATCGTGATCT 1213
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3881 AGAGCAGCATGACCAAGATCCTGAGCCCTTCCGCGCCCGCAACCCCGCCAGATCGTGATCT 3940

Qy 1214 ACCAGGCCCCCCTTGTACGTGGGCAGCGACCGCACCGTGGGGGCCAAGATCG 1273
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3941 ACCAGGCCCCCCTTGTACGTGGGCAGCGACCGCACCGTGGGGCCAAGATCG 4000

Qy 1274 AGGAGCTGCGCGAAGCACCTGCTGCGCGTGGCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGA 1333
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4001 AGGAGCTGCGCGAAGCACCTGCTGCGCGTGGCGCTGGGGCTTCACCACCCCGACAAGAAGCACCAGA 4060

Qy 1334 AGCTGCCCGGGGAAGGGGAGCTGGAGAGAGCTGGTGGGTGGGGCCCATCGGA 1393
         |||||||||||||||||||||||||||||||||||||||||
Db 4061 AGCTGCGCGAAGGGGAGCTGGAGAGAGCTGGTGGGTGGGCCCATCG 4120

Qy 1394 AGCTGCCCGGGGAAGGGGAGCTGGAACGACACTCAGAGCTGGTGGGTGGGCCCATCTGA 1453
         ||||||||||||||||||||||||||||||||||||||||||||
Db 4121 AGCTGCGCGGAAGGGGAGCTGGAACGACACTCAGAGCTGGTGGGTGGGCAAGCTGA 4180

Qy 1454 ACTGGGCCCAGATCTCACCCCCGGCATCAAGGTGCGCGAGCTGTGCAAGCTGTCTGCGCG 1513
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4181 ACTGGGCCCAGATCTCACCCCCGGCATCAAGGTGCGCGAGCTGTGCAAGCTGTCTGCGCG 4240

Qy 1514 GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGCCG 1573
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4241 GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGCCG 4300

Db 4301 AGAACCCGGAGATCCTGCGCGGCGAGCCCGTGCACGGCGGCGTGTACTACGACCCCAGCAAGGACC 4360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1634 TGGTGGCCGAGATCCAGAAGCAGGAGGGCCACGACCAGTGACCTACCAGGAGC 1693
         |||||||||||||||||||||||||||||||||||||
Db 4361 TGGTGGCCGAGATCCAGAAGCAGGAGGGCCACGACCAGTGACCTACCAGGAGC 4420

Qy 1694 CCTTCAAGAACCTGAAGACCGCAAGTACGCCAAGATGGCACCGGCCACCACGACG 1753
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4421 CCTTCAAGAACCTGAAGACCGCAAGTACGCCAAGATGGCACCGGCCACCACGACG 4480

Qy 1754 TGAAGCAGCTGACCAGGCCCGTGCAGAAGATCGCCATGAGAGATCGTGTAGTCTGGGGCA 1813
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4481 TGAAGCAGCTGACCAGGCCCGTGCAGAAGATCGCCATGAGAGATCGTGTAGTCTGGGGCA 4540

Qy 1814 AGACCCCAAGTTCCGCTGCCATCCAGAAGGAGACCTGGAGACCTGGAGACCGTGGACCGTGGACCGACT 1873
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4541 AGACCCCAAGTTCCGCTGCCATCCAGAAGGAGACCTGGAGACCTGGAGACCGTGGACCGTGGACCGACT 4600

Qy 1874 ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGACCACCCCCCCTGGTGAAGC 1933
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4601 ACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGACCACCCCCCCTGGTGAAGC 4660

Qy 1934 TGTGGTACCAGCTGGAAGAGCCCATCGTGGGCGTGGCAGCCGGCGACCTTCTACGTGACGGCG 1993
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4661 TGTGGTACCAGCTGGAAGAGCCCATCATCGGGCGTGGCAGCCGGCGACCTTCTACGTGGACGGCG 4720

Qy 1994 CCGCCAACCGCGAGACCAAGATCGGCAAGGCGTACGTGACCGACCGGCCGCAGA 2053
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4721 CCGCCAACCGCGAGACCAAGATCGGCAAGGCGTACGTGACCGACCGGCCGAGA 4780

Qy 2054 AGATCGTGAGCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 2113
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4781 AGATCGTGAGCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG 4840

Qy 2114 CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 2173
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4841 CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 4900

Qy 2174 TCATCCAGGCCCAGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCGAGCAGC 2233
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4901 TCATCCAGGCCCAGCCCGACAAGAGCGAGCTGGTGAACCAGATCATCGAGCAGC 4960

Qy 2234 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 2293
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4961 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 5020

Qy 2294 ACGAGCAGATCGACAAGCTGGTGAGCGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 2353
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5021 ACGAGCAGATCGACAAGCTGGTGAGCGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 5080

Qy 2354 TCGATGCGGCCATCGTGATCTACCAGTACCACGACCTGTACGTGGGCAGCGGCC 2413
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 5081 TCGATGCGGCCATCGTGATCTACCAGTACCACGACCTGTACGTGGGCAGCGGCC 5140

Qy 2414 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
         ||||||||||||||||||||||||||||||||||
Db 5141 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 5178

RESULT 9
ADC13279
ID   ADC13279
XX
AC   ADC13279;
XX
DT   18-DEC-2003 (first entry)
XX
DE   DNA of HIV construct TatRevNefgagCpolina_C SEQ ID NO 58.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;

XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316680P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 55; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intradermally, transdermally, intranasally,
CC   subcutaneously, intramucosally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 5184 BP; 1139 A; 1852 C; 1610 G; 583 T; 0 U; 0 Other;

Query Match   99.1%; Score 2434.8; DB 9; Length 5184;
Best Local Similarity 99.9%; Pred. No. 3.4e-293;
Matches 2436; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   14 TGGCCGAGGCCATGAGCGCCAGGCCACCGGACGCCCAGGCGCCAACATCCTGATGCAGCGCGAGCAACTTCA   73
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2741 TCGCCGAGGCCATGAGCGCCAGGCCACCGGACGCCCAGGCGCCAACATCCTGATGCAGCGCGAGCAACTTCA 2800

Qy   74 AGGGCCCCAAGGCCGCCATCATCATCGGCCGCAAGAAGGGCTGTGGAAGTGCGGCAAGGTGGGCCACCAGATCG 133
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2801 AGGGCCCCAAGGCCGCCATCATCATCGGCCGCAAGAAGGGCTGTGGAAGTGCGGCAAGGTGGGCCACCAGATCG 2860

Qy   134 ACTGCGCGCCCCCCGCAAGAAGGGCTGTGGAAGTGCGGCAAGGTGGGCCACCAGATCG 193
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2861 ACTGCGCGCCCCCCGCAAGAAGGGCTGTGGAAGTGCGGCAAGGTGGGCCACCAGATGA 2920

Qy   194 AGGACTGACCGAGGGCCAGGCCAGGCCAACTTCTTCCGCGAGACCTGGCCTTCCCCCCAGGGCA 253
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2921 AGGACTGACCGAGGGCCAGGCCAGGCCAACTTCTTCCGCGAGACCTGGCCTTCCCCCCAGGGCA 2980

Qy   254 AGGACCGCGAGTTCCCCAGCGAGACGAGAACCGCGGCCAACAGCCCCACCACCGCGGCGAGCTGC 313
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2981 AGGAGCCGCGAGTTCCCCAGCGAGACGAGAACCGCGCCAACAGCCCCACCACCGCGGCGAGCTGC 3040

Qy   314 AGTGCGCGGCGACAACCCCGCGAGGCGAGGCCCCAGCGAGGCCGCAGGCACCCTGAAACT 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3041 AGGTGCGCGGCGACAACCCCGCGAGGCGAGGCCCCAGCGAGGCCGCAGGCACCCTGAACT 3100

Qy   374 TCCCCCAGATCACCCTGTGGCAGGCCCCCCTGGCATCAAGGTGCGGCGGCCAGATCA 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3101 TCCCCCAGATCACCCTGTGGCAGGCCCCCCTGGCATCAAGGTGCGGCGGCCAGATCA 3160

Qy   434 AGGAGGCCCTGCTGGCCACCGGCGGCGCCGACGACCGTGCTGCTGGAGGAGATGAGCTGCCCG 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3161 AGGAGGCCCTGCTGGCCACCGGCGGCGCCGACGCACCGTGCTGCTGGAGGAGGATGAGCTGCCCG 3220

Qy   494 GCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGTGCGCCAGTACG 553
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3221 GCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGTGCGCCAGTACG 3280

Qy   554 ACCAGATCCTGATCGAGATCTGCGGCCACCGGCCAAGAAGGCCATCGGCCTGCTGATCGGCCCCA 613
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3281 ACCAGATCCTGATCGAGATCTGCGGCCACCGGCCAAGAAGGCCATCGGCCTGCTGATCGGCCCCA 3340

Qy   614 CCCCGTGAACATCATCGGCCGCAACATGTCGCCAGTGGGCTGCACCCTGAACTTCC 673
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3341 CCCCGTGAACATCATCGGCCGCAACATGTCGCCAGTGGGCTGCACCCTGAACTTCC 3400

Qy   674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCCATCTGCGAGGAGA 733
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCCATCTGCGAGGAGG 3460

Qy   734 TGAAGCAGTGCCCCTGACCGGAGAGAGAATCAAGGCCCTGACGGCCATCTGCGAGGAGA 793
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3461 TGAAGCAGTGCCCCTGACCGGAGAGAGAATCAAGGCCCTGACGGCCATCTGCGAGGAGA 3520

Qy   794 TGGGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGT 853
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3521 TGGGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCTACAACACCCCGTGT 3580

Qy   854 TCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGGCAGCTGA 913
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3581 TCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGGCAGCTGA 3640

Qy   914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCACCCCGGCCTGA 973
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3641 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGCATCCCCACCCCGGCCTGA 3700

Qy   974 AGAAGAAGAGAGCGGTGACCGTGCTGGACGTGGGCGACCGGCTACTTCAGCGTGCCCCTG 1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3701 AGAAGAAGAGAGCGGTGACCGTGCTGGACGTGGGCGACCGGCTACTTCAGCGTGCCCCTGG 3760

Qy   1034 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCGCATCAACAACGAGACCCCCG 1093
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3761 ACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACGAGACCCCCG 3820

Qy   1094 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCCCCAGCATCTTCC 1153
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3821 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCCCCAGCATCTTCC 3880

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCGAGATCGTGATCT 1213
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3881 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCGAGATCGTGATCT 3940

Qy   1214 ACCAGGGCCCCCTGTACGTGGGCAGGACCTGGGCCTTCACCACCCCGACCAAGAAGCACCAGATCG 1273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3941 ACCAGGGCCCCCTGTACGTGGGCAGGACCTGGGCCTTCACCACCCCGACCAAGAAGCACCAGA 4000

Qy   1274 AGGAGCTGCGCAAGCAGCACCTGCTGCGGGCTTCACCCCGACAGTGGACCTGGCACCCCATGG 1333
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4001 AGGAGCTGCGCAAGCAGCACCTGCTGCGGGCTTCACCCCGACAGTGGACCTGGCACCCCATGG 4060

Qy   1334 AGGAGCCCCCTTCCTGCTGCCATCGAGTCGACCTGAACGACATCCAGAAGCTGGGCAAGCTGC 1393
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4061 AGGAGCCCCCTTCCTGCTGCCATCGAGTCGACCTGAACGACATCCAGAAGCTGGGCAAGCTGC 4120

Qy   1394 AGTGCCGCGGAGAGAGCTGACCCTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 1453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4121 AGTGCCGCGGAGAGAGCTGACCCTGAACGACATCCAGAAGCTGGTGGGCAAGCTGA 4180

Qy   1454 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAGCTGTGCAAGCTGCTGCGCGG 1513
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4181 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGGCAGCTGTGCAAGCTGCTGCGCGG 4240

Qy   1514 GGGCCAAGGCCCTGACCGACGTCGTGCCGCTGACCGAGGAGGCCGAGCTGGAGCTGCCCG 1573
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 35; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g., a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 2457 BP; 564 A; 835 C; 758 G; 300 T; 0 U; 0 Other;

Query Match    98.8%;  Score 2428.6;  DB 7;  Length 2457;
Best Local Similarity   99.6%;  Pred. No. 2.2e-292;
Matches 2447;  Conservative 0;  Mismatches 4;  Indels 6;  Gaps 1;
```

```
Qy      1 GTCGACGGCCACCATGGCCGAGGCCATGAGCCAAGGCCAGCCCAACATCCTGATGCAG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTCGACGGCCACCATGGCCGAGGCCATGAGCCAAGGCCAGCCCAACATCCTGATGCAG   60

Qy     61 CGCAGCAACTTCAAGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCAGCAACTTCAAGGGCCCCAAGGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120

Qy    121 CACATCGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CACATCGCCCGCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAG  180

Qy    181 GGCCACCAGATGAAGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGCGAGGACCTGGCC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGCCACCAGATGAAGGACTGCACCGAGCGGCAGGCCAACTTCTTCCGCGAGGACCTGGCC  240

Qy    241 TTCCCCAGGGCAAGGGCCCGAGTTCCCCAGCGAGCAGAACCGCGCCACCCCACC  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCCCCAGGGCAAGGGCCCGAGTTCCCCAGCGAGCAGAACCGCGCCACCCCACC  300

Qy    301 AGCCGGAGCTGCAGGTGCGGCGGCGACAACCCCCGCAACCCCCGCGAGCGCCGAGCGCCAG  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCCGGAGCTGCAGGTGCGGCGGCGACAACCCCCGCAACCCCCGCGAGCGCCGAGCGCCAG  360
```

```
Qy   1574 AGAACCGCGAGATCCTGCCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACC  1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4301 AGAACCGCGAGATCCTGCCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACC  4360

Qy   1634 TGGTGGCCGAGATCCAGAGACCAGGGCCAGGGCCACGACCAGTGGACCTACCAGAGTCCAGAGAGC  1693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4361 TGGTGGCCGAGATCCAGAGACCAGGGCCAGGGCCACGACCAGTGGACCTACCAGAGC  4420

Qy   1694 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCACACCAAGCACG  1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4421 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCACACCAAGCACG  4480

Qy   1754 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATGTGATCTGGGGCA  1813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4481 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGATCTGGGGCA  4540

Qy   1814 AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACT  1873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4541 AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACT  4600

Qy   1874 ACTGGCCAGGCCACCTGCTGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGGTGAAGC  1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4601 ACTGGCAGGCCACCTGCTGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGGTGAAGC  4660

Qy   1934 TGTGGTACCAGCTGAGAAGGAGGAGCCATCATCGCGCCAGAGACTTCTACGTGACGGCG  1993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4661 TGTGGTACCAGCTGAGAAGGAGGAGCCATCATCGCGCCAGAGACTTCTACGTGACGGCG  4720

Qy   1994 CCGCCAACCGCGAGACCAAGATGGCAAGGCCAGGCCGGTACGTGACCGACGCCGGAGA  2053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4721 CCGCCAACCGCGAGACCAAGATCGGCAAGGCCAAGGCCGGTACGTGACCGACGCCGGAGA  4780

Qy   2054 AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG  2113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4781 AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGG  4840

Qy   2114 CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA  2173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4841 CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA  4900

Qy   2174 TCATCCAGGCCCAGCCCGACAAGAGCGAAGAGCGAGCTGGTGAACCAGATCATCGAGCAGC  2233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4901 TCATCCAGGCCCCAGCCCGACAAGAGCGAAGAGCGAGCTGGTGAACCAGATCATCGAGCAGC  4960

Qy   2234 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA  2293
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4961 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA  5020

Qy   2294 ACGAGCAGATCGACAAGTCGGTGAGCAAGGGCATCGGCAAGGTGCTGTTCCTGGACGGCA  2353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5021 ACGAGCAGATCGACAAGTCGGTGAGCAAGGGCATCGGCAAGGTGCTGTTCCTGGACGGCA  5080

Qy   2354 TCGATGGCGGCAATCTGATGTGATCTACCAGTACGATGACGACCTGTACGTGGGCAGCGGGCGCC  2413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5081 TCGATGGCGGCAATCTGATGTGATCTACCAGTACGATGACGACCTGTACGTGGGCAGCGGGCGCC  5140

Qy   2414 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  2451
          ||||||||||||||||||||||||||||||||||||||
Db   5141 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT  5178
```

```
RESULT 10
ACA03547
ID   ACA03547 standard; DNA; 2457 BP.
XX
AC   ACA03547;
XX
DT   22-MAY-2003  (first entry)
XX
DE   Synthetic DNA encoding immunogenic HIV peptide #30.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
```

```
Qy   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTG 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCACCCTGAACTTCCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTG 420

Qy   421 GGCGGCCAGATCAAGGAGGCCCTGCTGCTGGACACCGGCGACGACACCGTGCTGGAGGAG 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GGCGGCCAGATCAAGGAGGCCCTGCTGCTGGACACCGGCGACGACACCGTGCTGGAGGAG 480

Qy   481 ATGAGGCTGCCCGGCCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 ATGAGGCTGCCCGGCCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 540

Qy   541 GTGCCCAGTACGACCAGATCGATCGATCTGCGCGCAAGAGGCCATCGGCACCGTG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 GTGCCCAGTACGACCAGATCGATCGATCTGCGCGCAAGAGGCCATCGGCACCGTG 600

Qy   601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCGAACATGCTGACCCAGCTGGGCTGC 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 CTGATCGGCCCCACCCCGTGAACATCATCGGCCGCGAACATGCTGACCCAGCTGGGCTGC 660

Qy   661 ACCCTGAACTTCCCCATCAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 ACCCTGAACTTCCCCATCAGCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG 720

Qy   721 GACGGCCAAGTGAAGCAGTGGAAGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCGCC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GACGGCCAAGTGAAGCAGTGGAAGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCGCC 780

Qy   781 ATCTGGAGAGATGGAAGCAGTGGAAGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 ATCTGGAGAGATGGAAGCAGTGGAAGAGGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC 840

Qy   841 AACACCCCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 AACACCCCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900

Qy   901 TTCCGGCAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTCCGGCAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960

Qy   961 CCCGCCCGCCTGAGAGAAGAAGAGCGGCTGACGTGCTGGACGTGGACGTGACGCCTACTTC 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 CCCGCCCGCCTGAGAGAAGAAGAGCGGCTGACGTGCTGGACGTGGACGTGACGCCTACTTC 1020

Qy  1021 AGCGTGCCCCTGGACGAGGACTTCCGGAAGTACACACCGCCTTCACCGCCCATCAGCATCAAC 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 AGCGTGCCCCTGGACGAGGACTTCCGGAAGTACACACCGCCTTCACCGCCCATCAGCATCAAC 1080

Qy  1081 AACGACCCCCGCCATCCGCATCCGCTACCACGTGCTGCCCCAGGGCTGCGAAGGGCAGC 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 AACGACCCCCGCCATCCGCATCCGCTACCACGTGCTGCCCCAGGGCTGCGAAGGGCAGC 1140

Qy  1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGCCCTTCCGCGCCCGCGCAACCCC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGCCCTTCCGCGCCCGCGCAACCCC 1200

Qy  1201 GAGATCGATCTACCACCAGGAGCGGGCAGCGACTGGGCGCGACTCGGATGGATCGGCAGCAC 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 GAGATCGATCTACCACCAGGAGCGGGCAGCGACTGGGCGCGACTCGGATGGATCGGCAGCAC 1260

Qy  1261 CGCGCCAAGATCGAGACTCGAGAGCTGCGGGCTTCACCACCCCCCGAC 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 CGCGCCAAGATCGAGACTCGAGAGCTGCGGGCTTCACCACCCCCCGAC 1320

Qy  1321 AAGAAGCACCAGAGGAGCCCCCCTTCCTTCCTGCCCAT------CGAGTCGCACCGGCAAG 1374
         ||||||||||||||||||||||||||||||||||||||      |||||||||||||||||
Db  1321 AAGAAGCACCAGAGGAGCCCCCCTTCCTTCCTGCCCAGCCCTCAGCAGTCGCACCGGCAAG 1380

Qy  1375 TGGACCGTGCAGCCCATCGAGCTGCCCCAGAAGGAAGGAGCTGACGTGAAGCAGACATCCAG 1434
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TGGACCGTGCAGCCCATCGAGCTGCCCCAGAAGGAAGGAGCTGACGTGAAGCAGACATCCAG 1440
```

```
Db  1441 AAGCTGGTGGTGGCAAGAGTCAACTGGGCCGCCCAAGCTGCGATCGCTGACCCGACATGCCGCGATGTGCCGCAGGAG 1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  1495 CTGTGCAAGCTGCTGCGCGGCGGCGGCGGCCCAAGGCCCCTGACCGACATGTGCCCCTGACCGCGTGACGAGGAG 1554
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CTGTGCAAGCTGCTGCGCGGCGGCGGCGGCCCAAGGCCCCTGACCGACATCGTGCCCCTGACCGCGTGACGAGGAG 1560

Qy  1555 GCCGAGCTGGAGCTGGCCGAGAACCGGCCGAGATCCGGCGAGAACCGGCCGAGATCCGGCGTGCACGCGTGACCTAC 1614
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GCCGAGCTGGAGCTGGCCGAGAACCGGCTGGCCGAGAACCGGCCGAGATCCGGCGTGCACGCGTGACCTAC 1620

Qy  1615 TACCACCCCCAGCAAGGACCTGGTGGCCGAGATCCAGAGCCAGAACAGGGCCACGACCAGCCGACC 1674
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 TACGACCCCCAGCAAGGACCTGGTGGCCGAGATCCAGAGCCAGAACAGGGCCACGACCAGCCGACC 1680

Qy  1675 TACCAGATCTACCAGGAGCCCTTCAAGAACTTGAAGACCGGCGGCGTGCAGAAGATCGCCATGGAGCGC 1734
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACCAGATCTACCAGGAGCCCTTCAAGAACTTGAAGACCGGCGGCGTGCAGAAGATCGCCAAGATCGC 1740

Qy  1735 ACCGCCACCACCACCCCCATCGAGACCGTGCAGCAGCTGGAGCTGGAGAAGATCGCAGAAGAGACCTGAG 1794
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ACCGCCACCACCACCCCCATCGAGACCGTGCAGCAGCTGGAGCTGGAGAAGATCGCAGAAGAGACCTGAG 1800

Qy  1795 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCCGGATCCCGGAGTGTCGTGAAC 1854
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCCGGATCGGGAGTGGAGTTCGTGAAC 1860

Qy  1855 GAGACCCTGGTGGACCGACTGGCAGGACCCCATCGGCGACCTGGTACCAGCTGTGGACCAGCTGTGGAACCCCGAG 1914
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 GAGACCCTGGTGGACGGACTACTGGCAGGACCCCATCGGCGACCTGGTACCAGCTGTGGACCAGCTGTGGAAC 1920

Qy  1915 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGCCCATCATCGGCGCCGGCTACTGG 1974
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGCCCATCATCGGCGCCGGCTACTGG 1980

Qy  1975 ACTTCTACGTGGACGGGCCGCCGGAGATCGTGACGAGCACCACCAGAGAACCGAG 2034
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ACTTCTACGTGGACGGGCCGCCGGAGATCGTGAGCGAGCACCACCAGAGAACCGAG 2040

Qy  2035 ACGGACCCGGGGCCATCCAGTCAGCTGCCCTGCCCGAGGTGAACATCGTGACCGAC 2094
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ACGGACCCGGGGCCATCCAGTCAGCTGCCCTGCCCGAGGTGAACATCGTGACCGAC 2100

Qy  2095 CTGCAGGCCATCCAGCTGCCCTGCCCCTGCAGGACGGCAGCGCGAGGTGAACATCGTGACCGAC 2154
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGCAGGCCATCCAGCTGCCCTGCCCCTGCAGGACGGCAGCGCGAGGTGAACATCGTGACCGAC 2160

Qy  2155 AGCCAGTACCCCGGCATCATCCAGCGCCTGTGCTGCCCAGGGCTGGAAGGGCAGC 2214
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 AGCCAGTACCCCGGCATCATCCAGCGCCTGTGCTGCCCAGGGCTGGAAGGGCAGC 2220

Qy  2215 AACCAGATCATCGAGCAGCTGATCGAAGAGGAGAGAAGTGTACCTGGAGTCTGGGTGCCCGCC 2274
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AACCAGATCATCGAGCAGCTGATCGAAGAGGAGAGAAGTGTACCTGGAGTCTGGGTGCCCGCC 2280

Qy  2275 CACAAGGGCATCGGCGCAACGAGCAGCGGCATCGTCTGAGCAAGCTGCTGTGAGCAAGCAGCAAG 2334
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACAAGGGCATCGGCGCAACGAGCAGCGGCATCGTCTGAGCAAGCTGCTGTGAGCAAGCAGCAAG 2340

Qy  2335 GTGCTGTTCCTGGACGGCGCATCGATGCGGCGCATCGTCGATCTACCAGTACCAGTACATGGAGACCTG 2394
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GTGCTGTTCCTGGACGGCGCATCGATGCGGCGCATCGTCGATCTGATCTACCAGTACCATGGACGACCTG 2400

Qy  2395 TACGTTGGGCAGCGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCCGT 2451
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 TACGTGGCAGCGCGGCGGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCCGT 2457
```

```
RESULT 11
ADC13265
ID     ADC13265 standard; DNA; 2457 BP.
XX
```

```
XX
DT   18-DEC-2003   (first entry)
XX
DE   DNA of HIV construct p2Pol-opt-YM_C SEQ ID NO 44.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 41; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 2457 BP; 564 A; 835 C; 758 G; 300 T; 0 U; 0 Other;

Query Match       98.8%;    Score 2428.6;   DB 9;   Length 2457;
Best Local Similarity  99.6%;    Pred. No. 2.2e-292;
Matches 2447;   Conservative   0;   Mismatches   4;   Indels   6;   Gaps   1;

Qy    1  GTCGACGCCACCATGGCCAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  GTCGACGCCACCATGGCCAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGATGCAG  60

Qy   61  CGGCAGCAACTTCAAGGGCCCCAAGCGCCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  CGGCAGCAACTTCAAGGGCCCCAAGCGCCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC  120

Qy  121  CACATCGCCCGCAACTGCCGCGCCGCGCCCCCCAAGAAGAGGGCTGCTGGAAGTGCGGCAAGGAG  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  CACATCGCCCGCAACTGCCGCGCCGCGCCCCCCAAGAAGAGGGCTGCTGGAAGTGCGGCAAGGAG  180

Qy  181  GGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  GGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCC  240

Qy  241  TTCCCCCAGGGCACGGCCAAGGCCCCGAGTTCCCCAGGACAGAACCCGGCCAACAGCCCCACC  300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TTCCCCCAGGGCACGGCCAAGGCCCCGAGTTCCCCAGGACAGAACCCGGCCAACAGCCCCACC  300

Qy  301  AGCGGCGAGCTGCAGGTGCGGGGCGGACAACCCCGCAGCCCCGAGCGCGCCGAGCGCCCAG  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  AGCGGCGAGCTGCAGGTGCGGGGCGGACAACCCCGCAGCCCCGAGCGCGCCGAGCGCCCAG  360

Qy  361  GGCACCCTGAACTTCCCCGAGATCACCCTGGTGGCAGCGCCCCCTGGTGGTGAGCATCAAGGTG  420
         |||||||||||||||||||||||||||||||||||||||| ||||||
Db  361  GGCACCCTGAACTTCCCCGAGATCACCCTGGTGGCAGCGCGCCCCTGGTGGTGAGCATCAAGGTG  420

Qy  421  GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCGGACGACACCGTGCTCGAGGAG  480
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  421  GGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGGCGCACCGTGCTCGAGGAG  480

Qy  481  ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG  540

Qy  541  GTGCGCCAGTACGACCAGATCCTGATGGAGATCTGCGGCAAGAAGGCCATCGGCACCGTG  600
         |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  541  GTGCGCCAGTACGACCAGATCCTGATGGAGATCTGCGGCAAGAAGGCCATCGGCACCACCGTG  600

Qy  601  CTGATGGCCCCACCCCGTGAACATCATCGGCCGCAACATCATCGGCCCGGCTGGGCTGC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CTGATGGCCCCACCCCGTGAACATCATCGGCCGCAACATCATCGGCCCGGCTGGGCTGC  660

Qy  661  ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  ACCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATG  720

Qy  721  GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  GACGGCCCCAAGGTGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC  780

Qy  781  ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTAC  840

Qy  841  AACACCCCGGTTCGCCATCAAGAAGAAGGACAGCAGCAAGTGGCGCAAGCTGGTGGAC  900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  AACACCCCGGTTCGCCATCAAGAAGAAGGACAGCAGCAAGTGGCGCAAGCTGGTGGAC  900

Qy  901  TTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCGCCCAC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TTCCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGCCATCGCCCAC  960

Qy  961  CCGCCCGGCCTGAAGGAGAGCGTGACCGTGCTGGGCGGCGACGCCGTGACTTC  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  CCGCCCGGCCTGAAGGAGAGCGTGACCGTGCTGGGCGGCGACGCCGTGACTTC  1020

Qy 1021  AGGTGCCCCTGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAAC  1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  AGGTGCCCCTGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAAC  1080

Qy 1081  AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  AACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC  1140

Qy 1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCC  1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAACCCC  1200

Qy 1201  GAGATCGATCTACCAGGCCCCCTGTACGTGGGCAGCGACCTGGCAGCTGGCCAGGAGCAC  1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  GAGATCGATCTACCAGGCCCCCTGTACGTGGGCAGCGACCTGGCAGCTGGCCAGGAGCAC  1260

Qy 1261  CGGCCAAGATCGAGGAGCTGCGCGCACTGGGGCTTCACCGGGACCCCGGAC  1320
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  CGGCCAAGATCGAGGAGCTGCGCGCACTGGGGCTTCACCGGGACCCCGGAC  1320

Qy 1321  AAGAAGCACCAGAAGGAGAGCCCCCT-----CGAGCTGCACCCCCGACAAG  1374
         |||||||||||||||||||||||||| ||||||||||||||||||||||
Db 1321  AAGAAGCACCAGAAGGAGAGCCCCCTTCCGTGATGGGCTACGAGCTGCACCCCCGACAAG  1380

Qy 1375  TGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGACGCTGGAACGAGACTCCAG  1434
         |||||||||||||||||||||||||||||||
Db 1375
```

```
ID   ABL39959 standard; DNA; 2469 BP.
XX
AC   ABL39959;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Synthetic construct PR975(+) SEQ ID NO:30.
XX
KW   Human immunodeficiency virus type C; antigenic HIV type C protein;
KW   immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW   immunostimulant; gene therapy; gene; ds.
XX
OS   Human immunodeficiency virus; type C.
OS   Synthetic.
XX
PN   WO20204493-A2.
XX
PD   17-JAN-2002.
XX
PF   05-JUL-2001; 2001WO-US021241.
XX
PR   05-JUL-2000; 2000US-00610313.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UVST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2002-154920/20.
XX
PT   New polynucleotides encoding antigenic HIV Type C polypeptides, useful in
PT   applications including DNA immunization or generation of packaging cell
PT   lines, particularly in gene therapy.
XX
PS   Claim 1; Fig 8; 233pp; English.
XX
CC   The present invention describes expression cassettes comprising a
CC   polynucleotide sequence encoding a polypeptide comprising immunogenic HIV
CC   type C polypeptides. The expression cassettes comprise any of the HIV
CC   type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef
CC   (I). (I) have immunostimulant activity and can be used in gene therapy.
CC   The HIV type C polynucleotides are useful in applications including DNA
CC   immunisation, generation of packaging cell lines, and production of HIV
CC   Type C proteins. The polynucleotides are particularly useful in gene
CC   therapy and DNA immunisation applications. ABL39942 to ABL40054 and
CC   ABB06204 to ABB06215 represent sequences used in the exemplification of
CC   the present invention
XX
SQ   Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 U; 0 Other;

Query Match             98.3%;  Score 2415.4;  DB 6;  Length 2469;
Best Local Similarity   99.3%;  Pred. No. 9.6e-291;
Matches 2451;  Conservative    0;  Mismatches   6;  Indels  12;  Gaps  2;

Qy        1 GTCGACGCGCCACCATGGCCGAGGCCATGAGCCAGCGCCACCAGCGCCAACATCCTGATGCAG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GTCGACGCGCCACCATGGCCGAGGCCATGAGCCAGCGCCACCAGCGCCAACATCCTGATGCAG 60

Qy       61 CGCAGCAACTTCAAGGGCCCCAAGCGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CGCAGCAACTTCAAGGGCCCCAAGCGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGC 120

Qy      121 CACATCGCCCGCAACTGCCGCGCCCCCGCCGGCCAGGGCGCTGCTGGAAGGGCGCGCAAGGAG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CACATCGCCCGCAACTGCCGCGCCCCCGCCGGCCAGGGCGCTGCTGGAAGGGCGCGCAAGGAG 180

Qy      181 GGCCACCAGATGAAGGACTGCACGGAGCCCAGCGCCAACTTCTTCCGCGGGACCTGGCC 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GGCCACCAGATGAAGGACTGCACGGAGCCCAGCGCCAACTTCTTCCGCGGGACCTGGCC 240

Qy      241 TTCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACAGAACCGCGCCAACAGCCCCACC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 TTCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGACAGAACCGCGCCAACAGCCCCACC 300
```

```
Db    1381 TGGACCGTGCAGCCATCGAGCTGCCCGAGAGGAGAGCTGGACCGTGAACGACGACATCCAG 1440
Qy    1435 AAGCTGGTGGGCAAGCTGAACTGGGCGCGGGCGCCAAGGCCCTGACCGACAGGTGCGCCAG 1494
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 AAGCTGGTGGGCAAGCTGAACTGGGCGCGGGCGCCAAGGCCCTGACCGACAGGTGCGCCAG 1500

Qy    1495 CTGTGCAAGTCTGCGCGGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1554
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 CTGTGCAAGTCTGCGCGGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAG 1560

Qy    1555 GCCGAGCTGGAGCTGGCCGAGAACCGCGCGAGATCCTGCGCAGACCCGTGCACGGCGTGTAC 1614
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGCGAGATCCTGCGCGAGACCCGTGCACGGCGTGTAC 1620

Qy    1615 TACGACCCCAGCAAGGACCTGTGTCGCGAGATCCAGAAGCAGGGCCAGGACCAGTGGACC 1674
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 TACGACCCCAGCAAGGACCTGTGTCGCGAGATCCAGAAGCAGGGCCAGGACCAGTGGACC 1680

Qy    1675 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACACGCCAAGATGCGC 1734
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACACGCCAAGATGCGC 1740

Qy    1735 ACCGCCCACACCAACCGACTGAAGAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1794
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 ACCGCCCACACCAACCGACTGAAGAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800

Qy    1795 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCGCTGCCATCCAGAAGGAGACCTGG 1854
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 AGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCGCTGCCATCCAGAAGGAGACCTGG 1860

Qy    1855 GAGACCTGGTGACCACTGCCAGGCCCACCTGGATCCCCGAGTTGGGAGTTCGTGAAC 1914
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GAGACCTGGTGACCACTGCCAGGCCCACCTGGATCCCCGAGTTGGGAGTTCGTGAAC 1920

Qy    1915 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1974
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1980

Qy    1975 ACCTTCTACGTGGACGGCGCGCGCCAACCGCGAGATCGGCGACCAAGATCGGCCAAGCGGTACGTG 2034
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 ACCTTCTACGTGGACGGCGCGCGCCAACCGCGAGATCGGCGACCAAGATCGGCCAAGCGGTACGTG 2040

Qy    2035 ACCGACCGGGGCGGCGAGAAGATCGTGAGCGAGGCCTGACCGAGACCACCCAGAAGACCGAG 2094
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 ACCGACCGGGGCGGCGAGAAGATCGTGAGCGAGGCCTGACCGAGACCACCCAGAAGACCGAG 2100

Qy    2095 CTGCAGGCCATCCAGTTGCCCTGCAGCAGCGGCGAGGTGAACATCGTGACCGAC 2154
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 CTGCAGGCCATCCAGTTGCCCTGCAGCAGCGGCGAGGTGAACATCGTGACCGAC 2160

Qy    2155 AGCCAGTACGCGCCCTGGGCCATCATCCAGGCCCCAGCCCCGACAAGAGCGAGCTGGTG 2214
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 AGCCAGTACGCGCCCTGGGCCATCATCCAGGCCCCAGCCCCGACAAGAGCGAGCTGGTG 2220

Qy    2215 AACCAGATCATCGAGCAGCTGATCAAGAAGGAAGCGTGTACCTGAGCTGGCCCGCC 2274
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAAGCGTGTACCTGAGCTGGCCCGCC 2280

Qy    2275 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAACGAGATCGTGTTCAGCGTGGACACCTGACAGTGGCATGGACGACCTG 2334
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 CACAAGGGCATCGGCGGCAACGAGCAGATCGACAACGAGATCGTGTTCAGCGTGGACACCTGACAGTGGCATGGACGACCTG 2340

Qy    2335 GTGCTGTTCCTGGACGGCGGCATCATGTGCTGGGCGCATCGTGATCTACCAGTACGGCTAGCACCTG 2394
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 GTGCTGTTCCTGGACGGCGGCATCATGTGCTGGGCGCATCGTGATCTACCAGTACGGCTAGCACCTG 2400

Qy    2395 TACGTGGGGCAGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 TACGTGGGGCAGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
```

```
RESULT 12
ABL39959
```

```
Qy  1369  GACAAGTGGACCGTGCGAGCTGCAGCCGCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC  1428
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  GACAAGTGGACCGTGCGAGCTGCAGCCGCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGAC  1440

Qy  1429  ATCCAGAAGCTGTGGCAAGCTGAACTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGGTG  1488
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  ATCCAGAAGCTGTGGCAAGCTGAACTGAACTGGGCCAGCCAGATCTACCCGGCATCAAGGTG  1500

Qy  1489  CGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACCGACATCGTGCCCCTGACC  1548
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  CGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCCCTGACACATCGTGCCCCTGACC  1560

Qy  1549  GAGGAGGCCGGAGCTGGAGCTGCCGGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGC  1608
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  GAGGAGGCCGGAGCTGGAGCTGCCGGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGC  1620

Qy  1609  GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGGCAAGATCCGAGATCCAGAACAGAGC  1668
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  GTGTACTACGACCCCAGCAAGGACCTGGTGGCCGGCAAGATCCAGAACAGAGC  1680

Qy  1669  TGGACCTACCAGATCTACCAGGAGACCGGCAAGAACCTGAAGACCGGCAAGTACGCCAAG  1728
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  TGGACCTACCAGATCTACCAGGAGACCGGCAAGAACCTGAAGACCGGCAAGTACGCCAAG  1740

Qy  1729  ATGCGCACCGCCCACCACGTGAAGCAGCGTGACCGAGCCGGCCGTGCAGAAGATCGCC  1788
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  ATGCGCACCGCCCACCACGTGAAGCAGCGTGACCGAGCCGGCCGTGCAGAAGATCGCC  1800

Qy  1789  ATGGAGGCATCGTGATCTGGGCGCCAAGACCCCCAAGTTCGCCTGCCATCCAGAAGGAG  1848
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  ATGGAGGCATCGTGATCTGGGCGCCAAGACCCCCAAGTTCGCCTGCCATCCAGAAGGAG  1860

Qy  1849  ACCTGGGAGACCTGGTGACCACTACTGGCAGGCCCACCTGGATCCCGAGTGGGAGTTC  1908
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  ACCTGGGAGACCTGGTGACCACTACTGGCAGGCCCACCTGGATCCCGAGTGGGAGTTC  1920

Qy  1909  GTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGTGTGGAGAAGGACGCCCATCATCGGC  1968
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  GTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGTGTGGAGAAGGACGCCCATCATCGGC  1980

Qy  1969  GCCGAGACCTTCTACGTGACGGCGCGCAACCGGAGACCAAGATCGGCAAGAGGCCGGC  2028
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  GCCGAGACCTTCTACGTGACGGCGCGCAACCGGAGACCAAGATCGGCAAGAGGCCGGC  2040

Qy  2029  TACGTGACCGACCGGGGCCGGCAGAGATCGTGAGCGTGACCGAGACACCAGAGAAAG  2088
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  TACGTGACCGACCGGGGCCGGCAGAGATCGTGAGCGTGACCGACCACCAACCAGAAG  2100

Qy  2089  ACCGAGCTGCAGGCCATCCAGCTGGCCCTGGGCATCATCCAGCAGGCCGAGGTGAACATCGTG  2148
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  ACCGAGCTGCAGGCCATCCAGCTGGCCCTGGGCATCATCCAGCAGGCCGAGGTGAACATGTG  2160

Qy  2149  ACCGACAGCCAGTACGGCCAGTCCAGGCCCAGCAGAGAGCGGAGAGCAGAGAGCGAGAGCGAG  2208
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  ACCGACAGCCAGTACGGCCAGTCCAGGCCCAGCAGAGAGCGGAGAGCAGAGAGCGAGAGCGAG  2220

Qy  2209  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAAGAAGGTGTACCTGAGCTGGGTG  2268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  CTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAAGAAGGTGTACCTGAGCTGGGTG  2280

Qy  2269  CCCGCCCACAAGGGCATCGGCGCAACGAGCAGCTGGTGAGCAAGGGCATC  2328
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CCCGCCCACAAGGGCATCGGCGCAACGAGCAGCTGGTGAGCAAGGGCATC  2340

Qy  2329  CGCAAGGTGCGTGTTCCTGGACGGCATCGATGCGGCATCGTGATCTACCAGTAGTACATCGAC  2388
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  CGCAAGGTGCGTGTTCCTGGACGGCATCGATGCGGCATCGTGATCGTGATCTACCAGTACACATGGAC  2400

Qy  2389  GACCTGTACGTGGGCAGCGGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGGGGCTAGCACC  2448
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  GACCTGTACGTGGGCAGCGGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGGGGCTAGCACC  2460

Qy  2449  GGTGAATTC  2457
```

```
Qy  301  AGCCGGAGCTGCAGGTGCGCGGCGACAACACCCCGCGAGCGCCGGCGCCAG  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  AGCCGGAGCTGCAGGTGCGCGGCGACAACACCCCGCGAGCGCCGGCGCCAG  360

Qy  361  GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTG  420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GGCACCCTGAACTTCCCCAGATCACCCTGTGGCAGCGCCCCTGGTGAGCATCAAGGTG  420

Qy  421  GGGCGGCCAGATCAAGGAGGAGCCCCTTGCTGGACACCGCCGGCGACCTGTGGAGGAG  480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GGGCGGCCAGATCAAGGAGGAGCCCCTGCTGGACGACGACGACACCGCTGGAGGAG  480

Qy  481  ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCGCTTCATCAAG  540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  ATGAGCCTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCCTTCATCAAG  540

Qy  541  GTGCGCAGTACGACCCGTGATCGAGATCCTGATCGAGATCGGCGCGCAAGGCCATCGGCACCGTG  600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  GTGCGCAGTACGACCCGTGATCGAGATCCTGATCGAGATCGGCGCGCAAGGCCATCGGCACCGTG  600

Qy  601  CTGATCGAACTTCCCCATCAGCCCGTGACCCGTGAAGCTGAAGCCCGGCGGGCTGC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  CTGATCGGACCCCCATCAGCCCGTGACCCGTGAAGCTGAAGCCCGGCTGGGCTGC  660

Qy  661  ACCCTGAAGGTGAAGCAGTGTGATCCCCATCCGAGATCAGGACACCAAGGTGGTGAC  720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  ACCCTGAAGGTGAAGCAGTGTGATGATCCCCATCCGAGATCAGGACACCAAGGTGGTGAC  720

Qy  721  GACGGCCCAAGGTGAAGCAGCAGTGTGACCGAGAGAGTACCAAGGCGCACCCCTGACCGCC  780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  GACGGCCCAAGGTGAAGCAGCAGTGTGACCGAGAGAGTACCAAGGCCCGGCCCTGACCGCC  780

Qy  781  ATCTGCGAGAGATGAGAAGGAGGCAAGATCACCCAGATCAGGACATCGGCGCCCCCTAC  840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  ATCTGCGAGAGATGAGGAGGAGGCAAGATCACCCAGATCAGGACATCGGCCCCCTAC  840

Qy  841  AACACCCCGTGTTGCCATCAAGAAGGACGCACCAAGTGCGCAAGCTGGTGGAC  900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  AACACCCCGTGTTGCCATCAAGAAGGACGCACCAAGTGCGCAAGCTGGTGAC  900

Qy  901  TTCCGGCGAGCTGAACAAGCCCACCTTCAGATGACTTCTGGCAGCTGGAGGTGCAGCTGGGCATCCCCCAC  960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TTCCGGCGAGCTGAACAAGCCCACCTTCAGATGACTTCTGGCAGCTGGAGGTGCAGCTGGGCATCCCCCAC  960

Qy  961  CCCGCCGGCCTGAGAGATGAGAGAGCTTCCGGACTACACCGCCTTCACCGCCTTCCGGAGGGCTTCACCCTACTTC  1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  CCCGCCGGCCTGAGAGATGAGAGAGCTTCCGGACTACACCGCCTTCACCGCCTTCCGGAGGGCTTCACCTACTTC  1020

Qy  1021  AGCGTGCCCGCCGGCCCGGCGATCCGGACTACAACCGCCTTCCGCCAGATCAAC  1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021  AGCGTGCCCGCCGGCCCGGCGATCCGGACTACAACCGCCTTCCACCAGATCAAC  1080

Qy  1081  AACGAGACCCCGCCAGTACGGCGCATCCGGTGCTGCCCCAGGGCTGGAGGGGCAGC  1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081  AACGAGACCCCGCCAGTACAGCGCATCCGGTGCTGCCCAGGGCTGGAGGGGCAGC  1140

Qy  1141  CCCAGCCATCTTCCAGAGCAGCAGATCCGGTACGTGACGACCTGGAGATCGGC  1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141  CCCAGCCATCTTCCAGAGCAGCAGATCCGGTACGTGACGACCTGGAGATCGGC  1200

Qy  1201  GAGATCGTGATCTACCA------GGCCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGC  1254
          ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db  1201  GAGATCGTGATCTACCAGAGCAGCAGATCCGGTACGTGGGCAGCGACCTGGAGATCGGC  1260

Qy  1255  CAGCAGACCCGCCAAGTACGGCGCGAAGCTGCGGGGCTTCACCACC  1314
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  CAGCAGACCCGCCAAGTACGGCGCGAAGCTGCGGGGCTTCACCACC  1320

Qy  1315  CCCGACCAAGAAGCACCAGAGAGGAGCCCCTTCCT------CGAGCTGCACCCC  1368
          ||||||||||||||||||||||||||||||||||||      |||||||||||
Db  1321  CCCGACCAAGAAGCACCAGAGAGGAGCCCCTTCCTGTGATGGCTACGAGCTGCACCCC  1380
```

```
Db     2461 GGTGAATTC 2469
            |||||||||

RESULT 13
ADC13234
ID   ADC13234 standard; DNA; 3531 BP.
XX
AC   ADC13234;
XX
DT   18-DEC-2003  (first entry)
XX
DE   DNA of HIV construct GagPolmut_C SEQ ID NO 13.
XX
KW   expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW   Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS   Human immunodeficiency virus.
XX
PN   WO2003004620-A2.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021420.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (UYST-) UNIV STELLENBOSCH.
XX
PI   Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR   WPI; 2003-221593/21.
XX
PT   New expression cassette comprising a polynucleotide sequence encoding a
PT   polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT   Prot, or Rev polypeptide, useful for immunization, or generating
PT   packaging cell lines.
XX
PS   Disclosure; Fig 10; 301pp; English.
XX
CC   The invention relates to a novel expression cassette comprising a
CC   polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC   Int, Nef, p15RnaseH, Pol, Tat, Prot, or Rev polypeptide. The novel
CC   expression cassette can be used to treat HIV type C by gene therapy or
CC   used in the development of a vaccine. The gene delivery vector is
CC   administered intramuscularly, intramucosally, intranasally,
CC   subcutaneously, intradermally, transdermally, intravaginally,
CC   intrarectally, orally or intravenously. The expression cassette is useful
CC   for immunisation, generating packaging cell lines and producing HIV
CC   polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC   Type C related sequence of the invention.
XX
SQ   Sequence 3531 BP; 802 A; 1210 C; 1096 G; 423 T; 0 U; 0 Other;
```

```
Query Match     97.9%;  Score 2404.4;  DB 9;  Length 3531;
Best Local Similarity  99.1%;  Pred. No. 2.1e-289;
Matches 2417  Conservative  0;  Mismatches  21;  Indels  0;  Gaps  0;

Qy     14 TGGCCGAGGCCATGAGCGCCAGGCCACCAGGCCCAACATCCTGATGCAGGCGCAGCAACTTCA 73
          ||||||||||||| |||||||||||| || || |||||| ||||||||||||||||| | ||
Db   1088 TGGCCGAGGCCGATGAGCGCCAGGCCAACACCAGGCCCAACCGTGATGATGCAGAAGCAACTTTAAAA 1147

Qy     74 AGGGGCCCCAAGGCCCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGGCCACATCGCCCGCA 133
          |||||||||||| ||||||||||||| || ||||| ||||||||||||||||||||||||
Db   1148 ACTTGCGGCCCCCCCCCGCAAGGAGGGGCCAAGGAGGGGCCAAGGAGGGGCCACATCGCCCGCA 1207

Qy    134 ACTTGCGGCCCCCCCCCGCAAGGAGGAAGGGGCTGCTGGAAGTGCTGCTGGAAGTGGGGTGCTGCTGGAAGTGA 193
          
Qy    194 AGGACTGCACGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA 253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1268 AGGACTGCACGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA 1327

Qy    254 AGGCCCGGCGAGTTCCCCGACGAGCAGCAGAACCGGGCCAACAGCCCGCGCCGCGAGTGC 313
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db   1328 AGGCCCGGCGAGTTCCCCGACGAGCAGCAGAACCGGGCCAACAGCCCCGCGCCGCGAGTGC 1387

Qy    314 AGGTGCGGCGGCGACAACCCCGCAGGCGAGGAGGCCGGCGAGGCGCCGAGCGCCACCCTGAACT 373
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1388 AGGTGCGGCGGCGACAACCCCGCAGGCGAGGAGGCCGGCGAGGCGCCGAGCGCCACCCTGAACT 1447

Qy    374 TCCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTGGGCGGCCAGATCA 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1448 TCCCCCAGATCACCCTGTGGCAGCGCCCCTGTGAGCATCAAGGTGGGCGGCCAGATCA 1507

Qy    434 AGGAGGCCCTGCTGGACACCGGCCGACGACACCGTGCTGGAGGAGGATGAGCCTGCCCG 493
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1508 AGGAGGCCCTGCTGGACACCGGCCGACGACACCGTGCTGGAGGAGGATGAGCCTGCCCG 1567

Qy    494 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 553
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1568 GCAAGTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACG 1627

Qy    554 ACCAGATCTGATCGAGATCTGACCGGCAAGAAGCCATCGGCACCGTGCTGTGATCGGCCCCA 613
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   1628 ACCAGATCTGATCGAGATCTGACCGGCAAGAAGGCCATCGGCACCGTGCTGTGATCGGCCCCA 1687

Qy    614 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGGGCTCACCCTGAACTTCC 673
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1688 CCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGATCGGCTGGGCTCACCCTGAACTTCC 1747

Qy    674 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 733
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1748 CCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG 1807

Qy    734 TGAAGCAGTGGCCCCTGACCGAGGAAGATCAGGCCCTGACCGCCATCTGCGAGGAGA 793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1808 TGAAGCAGTGGCCCCTGACCGAGGAAGATCAGGCCCTGACCGCCATCTGCGAGGAGA 1867

Qy    794 TGGAGAAGGAGGGCCAAGATCACCACCAAGATCGGCCCCGAGAACCCCTACACACACCCCGTGT 853
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1868 TGGAGAAGGAGGGCCAAGATCACCACCAAGATCGGCCCCGAGAACCCCTACACACACCCCGTGT 1927

Qy    854 TCGCCATCAAGAAGAAGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGCGAGCTGA 913
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1928 TCGCCATCAAGAAGAAGACAGCACCAAGTGGCGCAAGTGGTGGACTTCCGCGAGCTGA 1987

Qy    914 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCCTGA 973
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1988 ACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCCACCCCGCCCTGA 2047

Qy    974 AGAAGAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCGCCTACTTCAGGGTGCCCCTGG 1033
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2048 AGAAGAGAGAGCGTGACCGTGCTGGACGTGGGCGACGCGCCTACTTCAGGGTGCCCCTGG 2107

Qy   1034 ACGAGGACTTCCGCAAGTACACCACCGCTTCACCATCCCAGCATCAACAACGAGACCCCCG 1093
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2108 ACGAGGACTTCCGCAAGTACACCACCGCTTCACCATCCCAGCATCAACAACGAGACCCCCG 2167

Qy   1094 GCATTCCGAGGCCTACCAGTACAACGTCGCTGCCCCAGTGCTGGAAGGGCAGCCCCAGCATCTTCC 1153
          |||| || || ||||||||||||||| |||||||||| |||||||||||||||||||||||||
Db   2168 GCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCC 2227

Qy   1154 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 1213
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2228 AGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCAGAACCCCGAGATCGTGATCT 2287

Qy   1214 ACCAGGCCCCCGTACGTGGGCAGCTGGAGATCGGCCAGAGATCGGCAAGGCGCCAAGATCG 1273
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2288 ACCAGGCCCCCGTACGTGGGCAGCTGGAGATCGGCCAGAGATCGGCAAGGCGCCAAGATCG 2347

Qy   1274 AGGAGTCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGA 1333
```

Db   3428 TCGATGCGGCATCGTGATCCTACCAGTACATGGAGAGACCTGTACGTGGGCAGCGCGGCC 3487
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2414 CTAGGATCGATTAAAAGCTTCCCGGGCTAGCACCGGT 2451
            ||||||||||||||||||||||||||||||
Db   3488 CTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCCGT 3525

RESULT 14
ACA03548
ID   ACA03548 standard; DNA; 2457 BP.
XX
AC   ACA03548;
XX
DT   22-MAY-2003 (first entry)
XX
DE   Synthetic DNA ecoding immunogenic HIV peptide #31.
XX
KW   Immunogenic HIV polypeptide; human immunodeficiency virus; HIV; vaccine;
KW   gene therapy; packaging cell line; humoral immune response;
KW   cellular immune response; gene delivery vector; DNA immunisation; ds.
XX
OS   Synthetic.
XX
PN   WO2003004657-A1.
XX
PD   16-JAN-2003.
XX
PF   05-JUL-2002; 2002WO-US021421.
XX
PR   05-JUL-2001; 2001US-0303192P.
PR   31-AUG-2001; 2001US-0316860P.
PR   16-JAN-2002; 2002US-0349728P.
PR   16-JAN-2002; 2002US-0349793P.
PR   16-JAN-2002; 2002US-0349871P.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Zur Megede J, Barnett SW, Lian Y;
XX
DR   WPI; 2003-221602/21.
XX
PT   New synthetic polynucleotides encoding antigenic HIV type B and/or type C
PT   polypeptides, useful as immunogenic compositions or vaccines for
PT   generating humoral or cellular immune responses against HIV in a subject,
PT   especially humans.
XX
PS   Example 1; Fig 36; 262pp; English.
XX
CC   The invention describes a synthetic polynucleotide encoding 2 or more
CC   immunogenic HIV polypeptides, where at least 2 of the polypeptides are
CC   derived from different HIV subtypes. The polynucleotide is useful for
CC   immunisation, generation of packaging cell lines, or production of HIV
CC   polypeptides. The polynucleotide and its encoded proteins are useful as
CC   immunogenic compositions or vaccines for generating humoral or cellular
CC   immune responses against HIV in a subject, or for inducing neutralising
CC   antibodies against HIV. The gene delivery vector comprising the
CC   polynucleotide is also useful for DNA immunisation of, or for generating
CC   an immune response (e.g. a humoral or cellular immune response) in, a
CC   subject such as a mammal, particularly a human. This sequence encodes a
CC   human immunodeficiency virus immunogenic peptide
XX
SQ   Sequence 2457 BP; 568 A; 830 C; 758 G; 301 T; 0 U; 0 Other;

Query Match   97.8%; Score 2403.4; DB 7; Length 2457;
Best Local Similarity   99.3%; Pred. No. 2.9e-289;
Matches 2439; Conservative   0; Mismatches 6; Indels 12; Gaps 2;

Qy   7 GCCACCATGCCAGGCCATGAGCGACCGCCAACATCCTGATGCAGCGCAGC 66
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GCCACCATGCCGAGGCCATGAGCGACCGCCACCAGCGCCAACATCCTGATGCAGCGCAGC 60

Qy   67 AACTTCAAGGGCCCCAAGGCGCATCATCAAGTGCTTCAACTGGCCGCAAGGAGGGCCACATC 126
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 2348 AGGAGCTGCGCAAGCAGCACCTGCTCGCGCTGGGGCTTCACCACCCCGACAAGAGCACCAGA 2407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 1334 AGGAGCCCCCTTCCTGCACCGGACGGTGCACCCGGACAAGTGGACCGTGCAGCCCATCG 1393
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2408 AGGAGCCCCCCTTCCTGCACCGGACGGTGCACCCGGACAAGTGGACCGTGCAGCCCATCG 2467

Qy 1394 AGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCGAGAGCTGTGGGCAAGCTGA 1453
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2468 AGCTGCCCGAGAAGGAGAGCTGGACCGTGAACGACATCGAGAGCTGTGGGCAAGCTGA 2527

Qy 1454 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCCG 1513
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2528 ACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCCG 2587

Qy 1514 GCGGCAAGGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGCTGGAGCTGGCCG 1573
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2588 GCGGCCAAGGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGCTGGAGCTGGCCG 2647

Qy 1574 AGAACCGCGGAGATCTCGCGGAGCCCGTGCGCGGCGGTGTACTACGACCCCAGCAAGGACC 1633
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2648 AGAACCGCGGAGATCTCGCGGAGCCCGTGCGCGGCGGTGTACTACGACCCCAGCAAGGACC 2707

Qy 1634 TGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTCGACCTACCAGAGC 1693
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 2708 TGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTCGACCTACCAGGAGC 2767

Qy 1694 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACCAACGACG 1753
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2768 CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACCAACGACG 2827

Qy 1754 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATGCCCATGGAGAGCATCGTGATCTGGGGCA 1813
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2828 TGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCAGAAGATCGTGATCTGGGGCA 2887

Qy 1814 AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGACCTGGTGGACCGACT 1873
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2888 AGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACCGACT 2947

Qy 1874 ACTGCGAGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGTGAAGC 1933
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2948 ACTGCGAGCCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCCCCCCCTGTGAAGC 3007

Qy 1934 TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCG 1993
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3008 TGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAGACCTTCTACGTGGACGGCG 3067

Qy 1994 CCGCCAACCGGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGCCGCAGA 2053
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3068 CCGCCAACCGGCGAGACCAAGATCGGCAAGGCCGGCTACGTGACCGACCGGGGCCGCAGA 3127

Qy 2054 AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTG 2113
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3128 AGATCGTGAGCCTGACCGAGACCACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTG 3187

Qy 2114 CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 2173
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3188 CCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCA 3247

Qy 2174 TCATCCAGGCCCAGCCCGACAAGAGCGAGTGGTGAACCAGATCATCGAGCAGC 2233
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3248 TCATCCAGGCCCCAGCCCGACAAGAGCGAGTGGTGAACCAGATCATCGAGCAGC 3307

Qy 2234 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 2293
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3308 TGATCAAGAAGGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCA 3367

Qy 2294 ACGAGCAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 2353
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3368 ACGAGCAGCAGATCGACACAGTCGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCA 3427

Qy 2354 TCGATGCGGCATCGTGATCCTACCAGTACATGGAGAGACCTGTACGTGGGCAGCGCGGCC 2413
          ||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    61  AACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATC  120

Qy   127  GCCGCCAACTGCCGCGCCGCCCCGCTGCTGGAAGTGCCGCAAGGAGGCCAC  186
          |||||||||||||||||||||||||||||||||||||||
Db   121  GCCGCCAACTGCGCGCCGCCCCGCCCAAGAAGGAGGCCTGCTGGAAGTGCCGCAAGGAGGAGGCCAC  180

Qy   187  CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC  246
          |||||||||||||||||||||||||||||||
Db   181  CAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCC  240

Qy   247  CAGGGCAAGGACCGCGGAGTTCCCAGCGAGCAGAACCGCGCCAACCAGCCCCACCAGCCGC  306
          |||||||||||||||||||||||||||
Db   241  CAGGGCAAGGACCGCGGAGTTCCCAGCGAGCAGAACCGCGCCAACCAGCCCCACCAGCCGC  300

Qy   307  GAGCTGCAGGTGCGCGGCGGACAACCCCGCAGCGAGGCCGGCGCCGAGGGCCAGGGCCACC  366
          |||||||||||||||||||||||||||
Db   301  GAGCTGCAGGTGCGCGGCGGGACAACCCCGCAGCGAGGCCGGCGCCGAGGGCCAGGGCCACC  360

Qy   367  CTGAACTTCCCCCAGATCACCCTGTGTGCAGCGCGCCCCCTGTGGTGAGCATCAAGTGGGCGGC  426
          |||||||||||||||||||||||||||
Db   361  CTGAACTTCCCCCAGATCACCCTGTGCAGCGCGCCCCCTGGTGAGCATCAAGTGGGCGGC  420

Qy   427  CAGATCAAGGAGCCCTGCTGACACCGGCCGCGACACCACCGTGCTGGAGGATGAGC  485
          |||||||||||||||||||||||||||
Db   421  CAGATCAAGGAGCCCTGCTGACACCGGCCGCGACACCACCGTGCTGGAGGATGAGC  480

Qy   487  CTGCCCCGGCAAGTGGAAGCCCAAGATCGGCGGCGCTTCATCAAGTGCGC  546
          |||||||||||||||||||||||||||
Db   481  CTGCCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCGCTTCATCAAGGTCGC  540

Qy   547  CAGTACGACCAGATCCTGATCGGCGCGAGAGGCCATCGACCGTGCTGCTGATC  606
          |||||||||||||||||||||||||||
Db   541  CAGTACGACCAGATCCTGATCGGCGCGAGAGGCCATCGACCGTGCTGCTGATC  600

Qy   607  GGCCCCACCCCCTGAACATCATCGGCGCGCAGCGTGGGCTGCCACCCTG  666
          |||||||||||||||||||||||||||
Db   601  GGCCCCACCCCCTGAACATCATCGGCGCGCAGCGTGGGCTGCCACCCTG  660

Qy   667  AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGCATGGACGGC  726
          |||||||||||||||||||||||||||
Db   661  AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCGCATGGACGGC  720

Qy   727  CCCAAGGTGAAGCAGTGCCCCTGACCGCCATCTGC  786
          |||||||||||||||||||||||||||
Db   721  CCCAAGGTGAAGCAGTGCCCCTGACCGCCATCTGC  780

Qy   787  GAGGAGATGGAAGGAGGAGGCAAGATCACCAGGAGCACCAGAACACC  846
          |||||||||||||||||||||||||||
Db   781  GAGGAGATGGAGAAGGAGGAGGCAAGATCACCAAGATCACCAGAACACC  840

Qy   847  CCCGTGTTCGCCATCAAGAGAGGAGGACTTCTGGGAGGTGCGCAAGCTGGGCATCCCCCCCC  906
          |||||||||||||||||||||||||||
Db   841  CCCGTGTTCGCCATCAAGAGAGGAGGACTTCTGGGAGGTGCGCAAGCTGGGCATCCCCCCGCC  900

Qy   907  GAGCTGAACAAGCCCACCCCAGAGCGTGGAGCTGGCAGCTTCCGC  966
          |||||||||||||||||||||||||||
Db   901  GAGCTGAACAAGCCCACCCCAGAGCTGGAGCTGGCAGCTTCCGC  960

Qy   967  GGCCTGGAAGAAGAAGAGAGAGCGTGACCGTGGGCGCAGCGTGCTCTGCGACGGTGGGCACGTG  1026
          |||||||||||||||||||||||||||
Db   961  GGCCTGGAAGAAGAAGAGAGAGCGTGACCGTGGGCGGCGACGTGGGCGACTTCAGCGTG  1020

Qy  1027  CCCCTGGACGAGGAGAGGACTTCCGCAAGTACAACGCCGCCTTCACCATCCCAGATCAACAACGAG  1086
          |||||||||||||||||||||||||||
Db  1021  CCCCTGGACGAGGAGAGGACTTCCGCAAGTACAACGCCGCCTTCACCATCCCAGATCAACAACGAG  1080

Qy  1087  ACCCCCGGCATCGGCATCCGTACCAGTACAACGTGCTGCCCTACCAGTCAAGGGCTGGAAGGGCAGCCCCAGC  1146
          |||||||||||||||||||||||||||
Db  1081  ACCCCCGGCATCGGCATCCGTACCAGTACAACGTGCTGCCCTACCAGTCAAGGGCTGGAAGGGCAGCCCCAGC  1140

Qy  1147  ATCTTCCAGAGCAGCAGCAATGACCAAGAATGATCGGCGAGGACCTTCGGAGCCCTTCGCGGAGCCCCAAGCCCCGAGATC  1206
          |||||||||||||||||||||||||||
Db  1141  ATCTTCCAGAGCAGCAGCATGACCAAGAATGATCGGCGAGACTCCTGGAGCCCTTCGCGGAGCCCCAACCCCGAGATC  1200
```

```
Qy  1207  GTGATCTACCA------GGCCCCCTGTACGTGGCAGTGGGCAGGACTCTGGAGATCGGCCAGCAC  1260
          |||||||||                |||||||||||||||||||||||||||||
Db  1201  GTGATCTACCAGTACATGGACGCATCGGAGCTGTACGTGGGCACGGACACCTGGAGAATCGGCCAGCAC  1260

Qy  1261  CGGCGCCAAGATGAGGAGCTCGGCGAGCAAGCACCTGCGTGCTGCCGGCTTCGCCGCTGGGCGCTGGGCGCCTTCACCACCCCCGAC  1320
          |||||||||||||||||||||||||||
Db  1261  CGGCGCCAAGATCGAGGAGCTCGGCGGCAAGCACCTGCCCGGCTTCGCCGCTGGGCGCTGGGCGCCTTCACCACCCCCGAC  1320

Qy  1321  AAGAAGCACCAGAAGGAGGCCCCTTCCTTGCCCAT------CAGCTCGACCCCGACAAG  1374
          ||||||||||||||||||||||||||||            ||||||||||||||
Db  1321  AAGAAGCACCAGAAGGAGGCCCCCTTCCTTGTGATGGGCTACGAGCTGCCCGACAAG  1380

Qy  1375  TGGACCGTGCAGCCCATCGAGCTCGCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG  1434
          |||||||||||||||||||||||||||
Db  1381  TGGACCGTGCAGCCCATCGAGCTCGCCGAGAAGGAGAGCTGGACCGTGAACGACATCCAG  1440

Qy  1435  AAGCTGGTGGGCGGCAAGCTGAACTGGGCGCGCAGCCAGATCTACCCCGGCATCAAGGTGTCGCCCAG  1494
          |||||||||||||||||||||||||||
Db  1441  AAGCTGGTGGGCGGCAAGCTGAACTGGGCGCGCAGCCAGATCTACCCCGGCATCAAGGTGTCGCCAG  1500

Qy  1495  CTGTGCAAGCTGCTGCGCGGGCGCAACCGGGAAGCCGGCGCCCGTGACCGCGGCGTGACGCGGCCGTGTAC  1554
          |||||||||||||||||||||||||||
Db  1501  CTGTGCAAGCTGCTGCGCGGGCGCAACCGGGAAGCCGGCGCCCGTGACCGCGGCGTGACGCGGCCGTGTAC  1560

Qy  1555  GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGGAGCCCGTGCGCGGGCCACCAGTGGACC  1614
          |||||||||||||||||||||||||||
Db  1561  GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGCGGAGCCCGTGCGCGGGCCACCAGTGGACC  1620

Qy  1615  TACGACCCAGCAAGGACCTGGTGGCCAAGCAAGCAGCAGGGCCGCGAAGTACGGCCAAGATGCGC  1674
          |||||||||||||||||||||||||||
Db  1621  TACGACCCCAGCAAGGACCTGGTGGCCAAGCAAGCAGCAGGGCCGCGAAGTACGGCCAAGAATGCGC  1680

Qy  1675  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGCTGAGAGCTGGAGCTGCGCAAGAGCGGCCAAGATGCGC  1734
          |||||||||||||||||||||||||||
Db  1681  TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGCTGAGAGCTGGAGCTGCGCAAGAGCGGCCAAGATGCGC  1740

Qy  1735  ACCGGCCCACACCAACGAGTGAAGCAGCTGACCGAGGCCGTGACCGAGAAGATCGGCCATGAGG  1794
          |||||||||||||||||||||||||||
Db  1741  ACCGCCCACACCAACGAGTGAAGCAGCTGACCGAGGCCGTGACCGAGAAGATCGGCCATGAGG  1800

Qy  1795  AGCATCGTGATCTGGGCGCAAGGCAGACCCCCAAGTTCCGCCTGCCATCCGAGAGGAGACCTGG  1854
          |||||||||||||||||||||||||||
Db  1801  AGCATCGTGATCTCGGGCGCAAGGCAGACCCCCAAGTTCCGCCTGCCATCCGAGAAGGAGACCTGG  1860

Qy  1855  GAGACCTGGTGGACCGACTACTGGCAGGCGCCAGCACTGGAGAAGGAGTTCGTGAAC  1914
          |||||||||||||||||||||||||||
Db  1861  GAGACCTGGTGGACCGACTACTGGCAGGCGCCACCTGGAGAAGGAGTTCGTGAAC  1920

Qy  1915  ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAAGGAGCCCATCATCGGCCGCCGAG  1974
          |||||||||||||||||||||||||||
Db  1921  ACCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAAGGAGACCAAGATCGGCCGCCGAG  1980

Qy  1975  ACCTTCAGTGGACGGCGCGCCATCATCGGCAAGGGCCGGCGGCCTACGTG  2034
          |||||||||||||||||||||||||||
Db  1981  ACCTTCAGTGGACGGCGCGCCATCATCGGCAAGGGCCGGCGGCCTACGTG  2040

Qy  2035  ACCGACGGGGCGCCGCAGAGATCTGAGCCTGACCGAGACACCAGAAGACCGAG  2094
          |||||||||||||||||||||||||||
Db  2041  ACCGACGGGGCGCCGCAGAGATCGTGAGCCTGACCGAGACACCACCAGAAGACCGAG  2100

Qy  2095  CTGCAGGCCATCAGCTGGCCCTGACAGCGGGGAGCTGAACATCGTGACCGAC  2154
          |||||||||||||||||||||||||||
Db  2101  CTGCAGGCCATCAGCTGGCCCTGACAGCGGGGAGCTGAACATCGTGACCGAC  2160

Qy  2155  AGCCAGTACGCGCCTGGGCATCATCCAGGGCGGTGGAAGGGCAGCCCCAGC  2214
          |||||||||||||||||||||||||||
Db  2161  AGCCAGTACGCGCCTGGGCATCATCCAGGGCGGTGGAAGGGCAGCCCCAGC  2220

Qy  2215  AACCAGATCATCGAGCAGCTGATCAAGAAGAAGGTGTACCTGAGCTGGTGGCGCGCC  2274
          |||||||||||||||||||||||||||
Db  2221  AACCAGATCATCGAGCAGCTGATCAAGAAGAAGGTGTACCTGAGCTGGTGGCGCGCC  2280
```

```
Qy    67 AACTTCAAGGGCCCCAAGCGCCATCATCAAGTGTTCAACTGCGCGGCAAGGAGGGCCACATC 126
         ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    61 AACTTCAAGGGCCCCAAGCGCCATCATCAAGTGCTTCAACTGCGCGGCAAGGAGGGCCACATC 120

Qy   127 GCCGCAACTGCGCGCGCCCCCCCGAAGAAGGGCGCTGCTGGAAGTGCGGGCAAGAGGAGGGCCAC 186
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GCCGCAACTGCGCGCGCCCCCCCGAAGAAGGGCGCTGGAAGTGCGGGCAAGAGGAGGGCCAC 180

Qy   187 CAGATGAAGGACTGCACCGAGGCGCCAGGCCAACTTCTTCCGGGAGGACTTGGCCTTCCCC 246
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CAGATGAAGGACTGCACCGAGGCGCCAGGCCAACTTCTTCCGGGAGGACTTGGCCTTCCCC 240

Qy   247 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGC 306
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAACCGCGCCAACAGCCCCACCAGCCGC 300

Qy   307 GAGCTGCAGTGCGGGGCAACCCCGCGCAGGCGCCGCCAGGGCAGGGCAGGGCCAC 366
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GAGCTGCAGTGCGGGGCAACCCCGCGCAGGCGCGCCAGGGCAGGGCAGGGCCAC 360

Qy   367 CTGAACTTCCCCCAGATCACCCTGGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCCGC 426
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CTGAACTTCCCCCAGATCACCCTGGTGGCAGCGCCCCCTGGTGAGCATCAAGGTGGGCCGC 420

Qy   427 CAGATCAAGGAGGCCCTGCTGGACACCCGGCGCGCCGACGACACCGGTGCTGGAGGATGAGC 486
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CAGATCAAGGAGGCCCTGCTGGACACCCGGCGCGCCGACGACACCGGTGCTGGAGGATGAGC 480

Qy   487 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGGTGCGC 546
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CTGCCCGGCAAGTGGAAGCCCAAGATGATCGGCGGCTTCATCAAGGTGCGC 540

Qy   547 CAGTACGACCAGATCCTGATCGAGAGGCCATCGGACCGTGCTGATC 606
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CAGTACGACCAGATCCTGATCGAGAGGCCATCGGACCGTGCTGATC 600

Qy   607 GGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGCCAGCTGGCTGCACCCTG 666
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 GGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGCCAGCTGGCTGCACCCTG 660

Qy   667 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCCGCATGGACGGC 726
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 AACTTCCCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCCGCATGGACGGC 720

Qy   727 CCCAAGGTGAAGCAGTAGCAGAGGCCCTGACCGAGGAAGATCAAGGCCCTGACCGGCCATCTGC 786
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 CCCAAGGTGAAGCAGTAGCAGAGGCCCTGACCGAGGAAGATCAAGGCCCTGACCGGCCATCTGC 780

Qy   787 GAGGAGATGGAAGAGGAGGCCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 846
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 GAGGAGATGGAAGAGGAGGCCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACACC 840

Qy   847 CCCGTGTTGCCATCAAGAAGAAGGACAGCAGCACCAAGTGCGCAAGCTGGTGGACTTCCGC 906
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 CCCGTGTTGCCATCAAGAAGAAGGACAGCAGCACCAAGTGCGCAAGCTGGTGGACTTCCGC 900

Qy   907 GAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 966
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 GAGCTGAACAAGGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCACCCCGCC 960

Qy   967 GGCCTGAGCAAGAAGAAGAGAGCGTGACCGTGGGCGACGTGGGCACGCGCCCTACTTCAGCGTG 1026
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCTGAGCAAGAAGAAGAGAGCGTGACCGTGGGCGACGTGGGCACGCGCCCTACTTCAGCGTG 1020

Qy  1027 CCCCTGGACGGAGCACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACAACAGCAG 1086
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 CCCCTGGACGGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCAACAACAACAGCAG 1080

Qy  1087 ACCCCGGCATCCGTACCAGTACAACGTGCTGCCCAGCATCCTGATGCAGCGCGCAGC 1146
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  2275 CACAAGGGCCATCGGCGGCCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCGATCCGCAAG 2334
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACAAGGGCCATCGGCGGCCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCGATCCGCAAG 2340

Qy  2335 GTGCTGTTCCTGGACGCGGCCATCGATGCGGCGCATCGTGATCTACCAGTACATGGACGACCTG 2394
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GTGCTGTTCCTGGACGCGGCCATCGATGCGGCGCATCGTGATCTACCAGTACATGGACGACCTG 2400

Qy  2395 TACGTGGCAGCGGCGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2451
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 TACGTGGCAGCGGCGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGT 2457
```

RESULT 15
ADC13266
ID ADC13266 standard; DNA; 2457 BP.
XX
AC ADC13266;
XX
DT 18-DEC-2003 (first entry)
XX
DE DNA of HIV construct p2Pol-opt_C SEQ ID NO 45.
XX
KW expression cassette; HIV Gag; Env; Int; Nef; p15RnaseH; Pol; Tat; Prot;
KW Rev; HIV type C; gene therapy; vaccine; immunisation; HIV; ds.
XX
OS Human immunodeficiency virus.
XX
PN WO200300620-A2.
XX
PD 16-JAN-2003.
XX
PF 05-JUL-2002; 2002WO-US021420.
XX
PR 05-JUL-2001; 2001US-0303192P.
PR 31-AUG-2001; 2001US-0316860P.
PR 16-JAN-2002; 2002US-0349871P.
XX
PA (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J, Barnett SW, Lian Y, Engelbrecht S, Van Rensburg EJ;
XX
DR WPI; 2003-221593/21.
XX
PT New expression cassette comprising a polynucleotide sequence encoding a
PT polypeptide including an HIV Gag, Env, Int, Nef, p15RnaseH, Pol, Tat,
PT Prot, or Rev polypeptide, useful for immunization, or generating
PT packaging cell lines.
XX
PS Disclosure; Fig 42; 301pp; English.
XX
CC The invention relates to a novel expression cassette comprising a
CC polynucleotide sequence encoding a polypeptide including an HIV Gag, Env,
CC Int, Nef, Pol, Tat, Prot, or Rev polypeptide. The novel
CC expression cassette can be used to treat HIV type C by gene therapy or
CC used in the development of a vaccine. The gene delivery vector is
CC administered intramuscularly, intramucosally, intranasally,
CC subcutaneously, intradermally, transdermally, intravaginally,
CC intrarectally, orally or intravenously. The expression cassette is useful
CC for immunisation, generating packaging cell lines and producing HIV
CC polypeptides. This polynucleotide sequence represents the DNA of an HIV
CC Type C related sequence of the invention.
XX
SQ Sequence 2457 BP; 568 A; 830 C; 758 G; 301 T; 0 U; 0 Other;

Query Match   97.8%; Score 2403.4; DB 9; Length 2457;
Best Local Similarity   99.3%; Pred. No. 2.9e-289;
Matches 2439; Conservative 0; Mismatches 0; Indels 12; Gaps 2;

```
Qy     7 GCCACCATGGCCGAGGCCATGAGCCTGGGCCAGGCCACCGCGCCAACATCCTGATGCAGCGCGCAGC 66
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2221 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGAGAAGGTGTACCTGAGCTGGTGCCCGCC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  2275 CACAAGGGCATCGGCGGCCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2334
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CACAAGGGCATCGGCGGCCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2340

Qy  2335 GTGCTGTTCCTGGACGGCCATCGATGGCCGGCATCGTGATCTACCAGTACTACGACACCTG 2394
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GTGCTGTTCCTGGACGGCCATCGATGGCCGGCATCGTGATCTACCAGTACTACGACACCTG 2400

Qy  2395 TACGTGGGCAGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGTAGCACCGGT 2451
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 TACGTGGGCAGCGGCGGCGCCCTAGGATCGATTAAAAGCTTCCCGGGGTAGCACCGGT 2457

Search completed: April 10, 2004, 07:32:35
Job time : 626.487 secs
```

```
Qy  1147 ATCTTCCAGAGCAGCATGACCAAGATCTGGAGCCCTTCCGCGCCCGCAACCCGAGATC 1206
         ||||||||||||||||||||  ||||||| |||||||||||||||||||||||||||
Db  1141 ATCTTCCAGAGCAGCATGATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACCCGAGATC 1200

Qy  1207 GTGATCTACCA------GGCCCCCTGTACGTGGGCAGCGACCTGGAGATCGGCCAGCAC 1260
         ||||||||||||       |||||||||||||||||||||||||||||||||||||||
Db  1201 GTGATCTACCAGTACATGAGCAGCTGTACATGGGCAGCGACCTGGAGATCGGCCAGCTC 1260

Qy  1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGTGGGCGCTTCACCACCCCGAC 1320
         |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGTGGGCTTCACCACCCCGAC 1320

Qy  1321 AAGAGCACCAGAAGGAGCCCCCTTCCTGCCCAT------CGAGCTGCACCCCGACAAG 1374
         |||||||||||||||||||||||||||||||       ||||||||||||||||||
Db  1321 AAGAGCACCAGAAGGAGCCCCCTTCCTGTGGATGGGCTACGAGCTGCACCCCGACAAG 1380

Qy  1375 TGGACCGTGCGAGCTGCGCCGAGGAGAGCTGGACCGTGGAACGACATCCAG 1434
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TGGACCGTGCAGCTGCGCCGAGGAGAGCTGGACCGTGGAACGACATCCAG 1440

Qy  1435 AAGCTGGTGGGCAAGCTGAACTGGGCCCAGATCTACCCCGCATCAAGGTGCGCCAG 1494
         |||||||||||||||||||||||||||||||  |||||||||||||||||||||||
Db  1441 AAGCTGGTGGGCAAGCTGAACTGGGCCACCGCCAGATCTACCCCGCATCAAGGTGCGCCAG 1500

Qy  1495 CTGTGCAAGCTGCTGCGCGGCGCGGCAAGGCCCTGACCGCGACATGTGCCCCTGACCGCGTGTAC 1554
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 CTGTGCAAGCTGCTGCGCGGCGCGGCAAGGCCCTGACCGCGACATGTGCCCCTGACCGCGTGTAC 1560

Qy  1555 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGGAACCGGCCCGTGCACCGGCGTGACC 1614
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 GCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGCGGAACCGGCCCGTGCACCGGCGTGACC 1620

Qy  1615 TACGACCCCAGCAAGGACCTGGTGCCGAGATCCAGAGCAGCAGGCCCACGACCAGTGACC 1674
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 TACGACCCCAGCAAGGACCTGGTGCCGAGATCCAGAGCAGCAGGCCCACGACCAGTGACC 1680

Qy  1675 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAAGATCGCGC 1734
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 TACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAAGATCGCGC 1740

Qy  1735 ACGGCCCACCAACGAGCAGTGAGCAGTCAGTGACCGAGAAGATCGCCATGGAG 1794
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ACGGCCCACCAACGAGCAGTGAGCAGTCAGTGACCGAGAAGATCGCCATGGAG 1800

Qy  1795 AGCATCGTGATCTGGGGCAAGACCCCCAGTTCCGCTGAGCCTGAGAAGATGGACCTGG 1854
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGCATCGTGATCTGGGGCAAGACCCCCAGTTCCGCTGAGCCTGAGAAGATGGACCTGG 1860

Qy  1855 GAGACCTGGTGGACCGACTACTGCAGGCCACCTGAGTGGGAGTTCGTGAAC 1914
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 GAGACCTGGTGGACCGACTACTGCAGGCCACCTGAGTGGGAGTTCGTGAAC 1920

Qy  1915 ACCCCCCGTGTGAAGCTGTGTACGAGGAAGGAGCCCATCATCGGCGCCGAG 1974
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 ACCCCCCGTGTGAAGCTGTGTACGAGGAAGGAGCCCATCATCGGCGCCGAG 1980

Qy  1975 ACCTTCTACGTGGACGCGGCGCCAACCCGAGATCGGCCAAGGCCGGCTACGTG 2034
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 ACCTTCTACGTGGACGCGGCGCCAACCCGAGATCGGCCAAGGCCGGCTACGTG 2040

Qy  2035 ACCGACCGGGCGGGCCAAGAAGATCGTGACCGAGCCTGACCAACCAGAAGACCCGAG 2094
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 ACCGACCGGGCGGGCCAAGAAGATCGTGACCGAGCCTGACCAACCAGAAGACCCGAG 2100

Qy  2095 CTGCAGGAGCCCATCCAGTCGCCCTGACCTGGCTCAGCGCCAGCAGCGAGTTGAACATCGTGACCGAC 2154
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGCAGGAGCCCATCCAGTCGCCCTGACCTGGCTCAGCGCCAGCAGCGAGTTGAACATCGTGACCGAC 2160

Qy  2155 AGCCAGTACGCCCTGGGCATCATCCAGCCGCCAGACAGCGAGCGCGTGTG 2214
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 AGCCAGTACGCCCTGGGCATCATCCAGCCGCCAGACAGCGAGCGCGTGTG 2220

Qy  2215 AACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGCCCGCC 2274
```